

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:46:24 ; Search time 86.5929 Seconds
(without alignments)
2086.045 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713
Sequence: 1 MAGIAAKLAKDREARAGLS.....PENTGTIEDLISWLCFSVL 700

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhcc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	99.3	700	6 Q9GLG1	Q9GLG1 macaca fasc
2	2568.5	69.2	705	13 Q9YIC1	Q9YIC1 colutrinx co
3	2480	66.8	724	13 Q8UW96	Q8UW96 xenopus lae
4	2464.5	66.4	714	6 Q9GLG2	Q9GLG2 macaca fasc
5	2463	66.3	715	13 Q42133	Q42133 gallus gall
6	2455.5	66.1	716	6 Q9N185	Q9N185 bos taurus
7	2453.5	66.1	716	6 Q9N0V6	Q9N0V6 bos taurus
8	2426	65.3	713	11 Q8C2J1	Q8C2J1 mus musculu
9	2381	64.1	704	13 Q918G2	Q918G2 brachydanto
10	2356.5	63.5	703	11 Q8BV9	Q8BV9 rattus norv
11	2335	62.9	530	11 Q8BV9	Q8BV9 mus musculu
12	2322.5	62.6	703	11 Q91VA3	Q91VA3 mus musculu
13	2309	62.2	702	13 Q918T0	Q918T0 xenopus lae
14	2279.5	61.4	650	6 Q9N0V7	Q9N0V7 mus scrofa
15	2142	57.7	421	11 Q8R486	Q8R486 mus musculu
16	2014	54.2	702	4 Q8N4R5	Q8N4R5 homo sapien

17	1959.5	52.8	709	6 Q9XSJ3	Q9XSJ3 oryctolagus
18	1954.5	52.6	709	6 Q9XSJ1	Q9XSJ1 bos taurus
19	1953.5	52.6	709	6 Q9XSJ2	Q9XSJ2 sus scrofa
20	1952.5	52.6	709	11 Q08702	Q08702 rattus norv
21	1946.5	52.4	701	13 Q8AYE8	Q8AYE8 xenopus lae
22	1943.5	52.3	709	11 Q88977	Q88977 mus musculu
23	1934.5	52.1	737	11 Q70376	Q70376 rattus norv
24	1925.5	51.9	757	11 Q70482	Q70482 rattus norv
25	1905.5	51.3	755	11 Q9QZF9	Q9QZF9 mus musculu
26	1882	50.7	690	4 Q14815	Q14815 homo sapien
27	1875.5	50.5	821	6 Q46596	Q46596 sus scrofa
28	1874.5	50.5	822	6 Q9TTH8	Q9TTH8 bos taurus
29	1867.5	50.3	822	6 Q9TTH8	Q9TTH8 ovis aries
30	1858.5	50.1	815	6 Q9GLG7	Q9GLG7 macaca fasc
31	1847	49.7	690	11 Q9D805	Q9D805 mus musculu
32	1844	49.7	688	11 Q35919	Q35919 mus musculu
33	1841	49.6	674	11 Q35920	Q35920 rattus norv
34	1795	48.3	664	4 Q9NS74	Q9NS74 homo sapien
35	1610.5	43.4	925	5 Q9VT65	Q9VT65 drosophila
36	1592.5	42.9	925	5 Q96454	Q96454 drosophila
37	1521	41.0	381	11 Q91UZ9	Q91UZ9 mus musculu
38	1453	39.1	720	11 Q9ER56	Q9ER56 mus musculu
39	1354	36.5	555	5 Q9V8U6	Q9V8U6 drosophila
40	1291	34.8	493	11 Q9R113	Q9R113 rattus norv
41	1213.5	32.7	760	5 Q45033	Q45033 schistosoma
42	1207.5	32.5	758	5 Q96072	Q96072 schistosoma
43	1194.5	32.2	758	5 Q96071	Q96071 schistosoma
44	1161	31.3	575	5 Q8MQV0	Q8MQV0 homarus ame
45	1120.5	30.2	462	11 Q9ER54	Q9ER54 mus musculu

ALIGNMENTS

RESULT 1

Q9GLG1 PRELIMINARY; PRT; 700 AA.
AC Q9GLG1;
DT 01-MAR-2001 (T-REMBLrel. 16, Created)
DT 01-MAR-2001 (T-REMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (T-REMBLrel. 23, Last annotation update)
DE Calpain 2.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Nakajima T., Fukiage C., Azuma M., Shearer T.R.;
RT "Calpain isoforms in the eye of monkey."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF284441; AAG22771.1; -.
DR HSSP; P04574; 1ALV.
DR MEROPS; C02.002; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease_C2.
DR InterPro; IPR000169; SHPoc_acsite.
DR Pfam; PFO1067; Calpain_III; 1.
DR Pfam; PFO0036; Calpain_III; 1.
DR Pfam; PFO0648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN_III; 1.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; CyPc; 1.
DR SMART; SM00054; Efn; 2.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
SQ SEQUENCE 700 AA; 80026 MW; DCEEL6214F05057C CRC64;

Query Match 99.3%; Score 3687; DB 6; Length 700;
Best Local Similarity 99.1%; Pred. No. 7.5e-272;
Matches 694; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAGIAAKLAKDREAEGLSHERAIKYLNQDYEAALNCELEAGTLFODPSPAIPALGF 60
 Db 1 MAGIAAKLVORREAAEGLSHERAIKYLNQDYEAALNCELEAGTLFODPSPAIPALGF 60
 Qy 61 KELGPSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTINEEI 120
 Db 61 KELGPSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTINEEI 120
 Qy 121 LARVPLNOSFOENAGIFHFQFQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 180
 Db 121 LARVPLNOSFOENAGIFHFQFQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 180
 Qy 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQXALQKSLGCG 240
 Db 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQXALQKSLGCG 240
 Qy 241 SIDITSAADSEALITFOKLVKHAASVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDN 300
 Db 241 SIDITSAADSEALITFOKLVKHAASVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDN 300
 Qy 301 CPSNNTIDPEERERLTRRHEDGEPFMSFSDPLRHYSRLICNLTPDILTSDTYKKMKLTK 360
 Db 301 CPSNNTIDPEERERLTRRHEDGEPFMSFSDPLRHYSRLICNLTPDILTSDTYKKMKLTK 360
 Qy 361 MDGMMRGSTAGGCNRYNPTFMNPOYLKLEBEDEDEDESGCTFLVGLIQKHRRQR 420
 Db 361 MDGMMRGSTAGGCNRYNPTFMNPOYLKLEBEDEDEDESGCTFLVGLIQKHRRQR 420
 Qy 421 KMGEDMHTIGFYIYVEPEELSGQTNHLSKNFPLTNRARSDFPINREVLNRPKLPBG 480
 Db 421 KMGEDMHTIGFYIYVEPEELSGQTNHLSKNFPLTNRARSDFPINREVLNRPKLPBG 480
 Qy 481 EYLIIVPSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEPDISSEDDIDGRRRLPAQ 540
 Db 481 EYLIIVPSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEPDISSEDDIDGRRRLPAQ 540
 Qy 541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 600
 Db 541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYIL 600
 Qy 601 WTQIOKQIYREIDVDSGTMSYEMKALEAGFMPQQLHOYVARFADQIILDFD 660
 Db 601 WTQIOKQIYREIDVDSGTMSYEMKALEAGFMPQQLHOYVARFADQIILDFD 660
 Qy 661 NFVRCIVLETLFKIFKQIDPENTGTIELDISWLCFSVL 700
 Db 661 NFVRCIVLETLFKIFKQIDPENTGTIELDISWLCFSVL 700

RESULT 2

Qy 09YIC1 PRELIMINARY; PRT; 705 AA.
 AC 09YIC1;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Quail calpain.
 GN OCLAL-1.
 OS Coturnix coturnix (Common quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID=9091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Maeda Y.;
 RT "cDNA sequence of quail skeletal muscle calpain.";
 DR EMBL; AB011080; BAA74564.1; -.
 DR HSSP; P04574; IALV.
 DR MEROPS; C02.003; -.

DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR000163; SHprot_acsite.
 DR Pfam; PF01067; Calpain_I; 1.
 DR Pfam; PF00036; efhand_2.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyapc_I.
 DR PROSITE; PS00225; CRYSTALLIN_BETA_GAMMA; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
 SQ SEQUENCE 705 AA; 8018 MW; 5C70634EB3AA93CD CRC64;

Query Match 69.2%; Score 2568.5; DB 13; Length 705;
 Best Local Similarity 65.9%; Pred. No. 1e-186;
 Matches 460; Conservative 116; Mismatches 121; Indels 1; Gaps 1;

Qy 3 GIAAKLAKDREAEGLSHERAIKYLNQDYEAALNCELEAGTLFODPSPAIPALGFKE 62
 Db 6 GIAARLORDLRAGEVGHNNAYKYLNQDYEAALNCELEAGTLFODPSPAIPALGFKE 65
 Qy 63 LGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTINEEI 122
 Db 66 LGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAALASLTINEEI 125
 Qy 123 RAVPLNOSFOENAGIFHFQFQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 182
 Db 126 RAVPHGGSFOEDYAGIFHFQFQYGEWVAVVDDRLPTDGLLPHSAGSEFWAL 185
 Qy 183 KAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQXALQKSLGCSI 242
 Db 186 KAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQXALQKSLGCSI 245
 Qy 243 DITSAADSEALITFOKLVKHAASVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDNCP 302
 Db 246 DITSAFMEAVITFOKLVKHAASVTGAEEVNSGSLQKLRIRNPMGEVETGRWMDNCP 305
 Qy 303 SMTIDPEERERLTRRHEDGEPFMSFSDPLRHYSRLICNLTPDILTSDTYKKMKLTKMD 362
 Db 306 EMDITDSEDEELQKKEDGEPFMSFSDPLRHYSRLICNLTPDILTSDTYKKMKLTKMD 365
 Qy 363 GMMRGSTAGGCNRYNPTFMNPOYLKLEBEDEDEDESGCTFLVGLIQKHRRQRKM 422
 Db 366 GMMRGSTAGGCNRYNPTFMNPOYLKLEBEDEDEDESGCTFLVGLIQKHRRQRKM 425
 Qy 423 GEDMHTIGFYIYVEPEELSGQTNHLSKNFPLTNRARSDFPINREVLNRPKLPBG 482
 Db 426 GEDMHTIGFYIYVEPEELSGQTNHLSKNFPLTNRARSDFPINREVLNRPKLPBG 485
 Qy 483 IIVPSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEPDISSEDDIDGRRRLPAQL 541
 Db 486 IIVPSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEPDISSEDDIDGRRRLPAQL 545
 Qy 542 AGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYILW 601
 Db 546 AGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIWMDLSDSGKLGKEFYILW 605
 Qy 602 TKIQKQIYREIDVDSGTMSYEMKALEAGFMPQQLHOYVARFADQIILDFD 661
 Db 606 NKIRSWLTIIRQVLDKSGTMSYEMKALEAGFMPQQLHOYVARFADQIILDFD 665
 Qy 662 FVRCIVLETLFKIFKQIDPENTGTIELDISWLCFSV 699
 Db 666 FVRCIVLETLFKIFKQIDPENTGTIELDISWLCFSV 703

RESULT 3

Qy 08UW96 PRELIMINARY; PRT; 724 AA.
 AC 08UW96;
 DT 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE mu/m-calpain large subunit.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miyoshi K., Sorimachi H., Tomioka S., Ishiura S., Suzuki K.;
 RT "Xenopus has a calpain most similar to mu/m-type chicken calpain.",
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB061521; BAB83262.1; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR000169; SHProl acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00054; Efh; 2.
 DR SMART; SM00054; Efh; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 SQ SEQUENCE 724 AA; 81170 MW; CA5CB090AE7A4288 CRC64;

Query Match 66.8%; Score 2480; DB 13; Length 724;
 Best Local Similarity 63.4%; Pred. No. 5.9e-180;
 Matches 455; Conservative 115; Mismatches 126; Indels 22; Gaps 3;

QY 3 GIAKLAKDREAAGLGSHERAIKYLNQDYALNCELEAGTLFODSPFAIPALGFKE 62
 6 GIASKLIKIDLKSGVSHCAVYQNDYSLKQCVESGILFEDNFAIPISILGFKE 65
 QY 63 LGFYSSKTRGMWRKPTLEICADPOFIIGATRTDICOAGLDCWLLAALSLTNEBILA 122
 66 LGFGSSKTRGVQWPRPSDIYDDPFIIGATRTDICOAGLDCWLLAALSLTNEBILH 125
 QY 123 RVPLNLSPOENYAGTFHPOFQWGEVWVVDRLPTKGELLFYHSAAGSEFWSLTE 182
 126 RVVPHGSGFQEDYAGTFHPOFQWGEVWVVDRLPTKGELLFYHSAAGSEFWSLTE 185
 QY 183 KAVAKINGCYEALSGATTEGFEPTGIAEWELKRPENLFFKIIKALOKSLGCSI 242
 186 KAVAKINGCYEALSGATTEGFEPTGIAEWELKRPENLFFKIIKALOKSLGCSI 245
 QY 243 DITSADSEAITTQKLVKGHAIVTGAEEVSNGLQKLRIRNPNWGEVWTRGNDNCP 302
 246 DITSADSEAITTQKLVKGHAIVTGAEEVSNGLQKLRIRNPNWGEVWTRGNDNCP 305
 QY 303 SMTTIDEEBERELTRRHEDSEFWMSFDFLRHYSRLICNLTPDTLSDTYKKKKTMD 362
 306 EMWEVDSEEDRLKMKEDSEFWMSFDFLRHYSRLICNLTPDTLSDTYKKKKTMD 365
 QY 363 GNNRSGSTAGCCRNYPNTFMNPOYLILKEEDEDSESGCTFLVGLIOKHRRROR 422
 366 GNNRSGSTAGCCRNYPNTFMNPOYLILKEEDEDSESGCTFLVGLIOKHRRROR 425
 QY 423 GEDMTTIGFQIYVEBEELSGQTNHLSKNPFLTNRARERDTPFLNREVLNRFPLPGEY 482
 426 GEDMTTIGFQIYVEBEELSGQTNHLSKNPFLTNRARERDTPFLNREVLNRFPLPGEY 485
 QY 483 ILVPSPEERKDDFCIRVSEKADYQAVDDLEALN-EEFDSSEDDIDGVRRLAQL 541
 486 ILVPSPEERKDDFCIRVSEKADYQAVDDLEALN-EEFDSSEDDIDGVRRLAQL 545
 QY 542 AGEDAIASAELOTLIRVLAKRODIDSGFSITCKIMVMDLSDSGKLGKEFYILW 601
 546 AGEDAIASAELOTLIRVLAKRODIDSGFSITCKIMVMDLSDSGKLGKEFYILW 605
 QY 602 TKIQ-----KYOKIYREIDVDRSGTWNSEYMRKALEAGFKMPCQ 641

DB 606 NKEFMAVMALLPACMGPVGINTPSLQTVFEREDLKSGLISSYEIRLAISSGYKLNNK 665
 QY 642 LHOIVARFADDDLIIDFNFRCLVRLTELFIPKOLDEPNTGTBLDLSMLCFSYV 699
 DB 666 LIQVLVARIADSDMGIDFDFVCCVLKLEAMFEPKALD-EGDGTAEMLNGEMLTMTW 722

RESULT 4

09GLG2 PRELIMINARY; PRT; 714 AA.

AC 09GLG2;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
 DE Calpain 1.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NC NCB1_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA Nakajima T., Fukiage C., Azuma M., Shearer T.R.;
 RT "Calpain isoforms in the eye of monkey."
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF284440; AAG22770.1; -.
 DR HSSP; P04574; 1ALV.
 DR MEROPS; C02.001; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR000169; SHProl acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; CyBpc; 1.
 DR SMART; SM00054; Efh; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 SQ SEQUENCE 714 AA; 81849 MW; 834690C14DE7AD8 CRC64;

Query Match 66.4%; Score 2464.5; DB 6; Length 714;
 Best Local Similarity 63.2%; Pred. No. 8.7e-179;
 Matches 440; Conservative 124; Mismatches 129; Indels 3; Gaps 2;

QY 3 GIAKLAKDREAAGLGSHERAIKYLNQDYALNCELEAGTLFODSPFAIPALGFKE 62
 13 GVSAGVQKQAKAKELGHRHNAIKYLGQDYEQRLARCLGSLTFRDEAFPPVQSLGFKD 72
 QY 63 LGFYSSKTRGMWRKPTLEICADPOFIIGATRTDICOAGLDCWLLAALSLTNEBILA 122
 73 LGFNSSKTYGIIKKRPTLELSNPQFIVDQATRTDICOAGLDCWLLAALSLTNEBILH 132
 QY 123 RVPLNLSPOENYAGTFHPOFQWGEVWVVDRLPTKGELLFYHSAAGSEFWSLTE 182
 133 RVVPHGSGFQNGYAGTFHPOFQWGEVWVVDRLPTKGELLFYHSAAGSEFWSLTE 192
 QY 183 KAVAKINGCYEALSGATTEGFEPTGIAEWELKRPENLFFKIIKALOKSLGCSI 242
 193 KAVAKINGCYEALSGATTEGFEPTGIAEWELKRPENLFFKIIKALOKSLGCSI 252
 QY 243 DITSADSEAITTQKLVKGHAIVTGAEEVSNGLQKLRIRNPNWGEVWTRGNDNCP 302
 253 DITSADSEAITTQKLVKGHAIVTGAEEVSNGLQKLRIRNPNWGEVWTRGNDNCP 312
 QY 303 SMTTIDEEBERELTRRHEDSEFWMSFDFLRHYSRLICNLTPDTLSDTYKKKKTMD 362
 313 SMTTIDEEBERELTRRHEDSEFWMSFDFLRHYSRLICNLTPDTLSDTYKKKKTMD 372
 QY 363 GNNRSGSTAGCCRNYPNTFMNPOYLILKEEDEDSESGCTFLVGLIOKHRRROR 420

```

Db 373 GTWRRGSTAGGCRNYPATFWNPPQKIRLDETDDDDDYDGRSGCSFVLTALMQKRRRR 432
421 KMGEDMHTIGTCIYEVEPEELSGQTNHLSKNFPLTNRARESDPFINREVLNFKLPG 480
433 RFGDMETIGTAVYEVPEELGQPLHKKRDFPLANRARESEOPINIREVSTRFLPG 492
481 EYLVPSFTFEPNKGDFCIRVSEKADQYAVDEIEANTL-EEFDISEDIDDDGRRLLFA 539
493 EYVVPSFTFEPNKGDFCIRVSEKADQYAVDEIEANTL-EEFDISEDIDDDGRRLLFA 552
540 QLAGEDAEISAFELQTIIRRLVLAQRDICKSDGSIETCKINVDMLDSDSGKLKKEFYI 599
553 QLAGEDAEISAFELQTIIRRLVLAQRDICKSDGSIETCKINVDMLDSDSGKLKKEFYI 612
600 LMTIKQKQKTYRELDVDRSGTMSYEMRKALBEAGFPMQOLQVIVARPADQIITDF 659
613 LMMIRINVLSTIFRRFDLDRSGMSAYEMRKALBEAGFPMQOLQVIVARPADQIITDF 672
660 DNFRCLVRLLETLFKIFKQLDPEMTGTIELDLISWL 695
673 DNFRCLVRLLETLFKIFKQLDPEMTGTIELDLISWL 708

```

RESULT 5

```

ID 042133 PRELIMINARY; PRT; 715 AA.
AC 042133;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE MUC1 protein (EC 3.4.22.11).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 1;
RP SEQUENCE FROM N.A.
RA Sorimachi H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN 2;
RP SEQUENCE OF 130-715 FROM N.A.
RX MEDLINE=95260862; PubMed=7742367;
RA Sorimachi H., Tsukahara T., Okada-Ban M., Sugita H., Ishiura S.,
RA Suzuki K.;
RT "Identification of a third ubiquitous calpain species--chicken muscle
RT expresses four distinct calpains.";
RL Biochim. Biophys. Acta 1261:381-393 (1995).
RN 3;
RP SEQUENCE FROM N.A.
RX MEDLINE=98129049; PubMed=9467868;
RA Jeong S.-Y., Sorimachi H., Lee H.-J., Ishiura S., Suzuki K.;
RT "Molecular cloning and characterization of cDNAs for the mu-type large
RT subunit and the small subunit of chicken calpain.";
RL Comp. Biochem. Physiol. 118B:539-547 (1997).
DR EMBL; AB007775; BAA22659.1; -.
DR HSSP; P04574; 1ALV.
DR MEROPS; CO2.001; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease_C2.
DR InterPro; IPR000169; SHProt_acsite.
DR Pfam; PF01067; Calpain_III.1.
DR Pfam; PF00036; ehand_3.1.
DR Pfam; PF00648; Peptidase_C2_1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III.1.
DR SMART; SM00230; Cyspc_1.
DR SMART; SM00054; EFh_2.
DR PROSITE; PS00018; EF_HAND_2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS_1.
DR SEQUENCE 715 AA; 8141 MW; 70679821C5E0AD6A CRC64;

```

Query Match 66.3%; Score 2463; DB 13; Length 715;
 Best Local Similarity 63.0%; Pred. No. 1,1e-178;
 Matches 439; Conservative 125; Mismatches 129; Indels 4; Gaps 2;

```

Qy 3 GIAKLAKDEAABEGLSHERAKIYNODYEARNCELEAGTLPQPSFPAISALGFE 62
13 GVSANOVORAKALIGQONAVRFGQDPAALRDCLRSGSLFRETFFPSASSLGFE 72
63 LGFYSKTRGMWRKRPFEICADPOFIIGATRTDIOGALGDCMLAALASLTNEIILA 122
73 LGFSSKTRGVTKRPFEICRHPQFIVDAKTRDIOGALGDCMLAALASLTNEIILA 132
123 RVPPLNGSFQENYAGIFHFQWQYGEWEVAVVDRLPTQKGLLPHASAGSEFWSALL 162
133 RVPVHGQSFQNGYAGIFHFQWQYGEWEVAVVDRLPTQKGLLPHASAGTEFWSALL 192
183 KAYAKINGCYEALSGATTTGGFDFPGIAEWELKPPNLEKTIQKALQKSLGCSI 242
193 KAYAKVNGCYEALSGSTSGFDFPGVTEWYDLRRPPADLYQIILKALERSLLGCSI 252
243 DITSAADEAITFOKLYGHAYSVTGAEEVNSGLQKLRIRNPMGEVETGRMNDNC 302
253 DITSAPFMEAVTTRKLVKGAHYSVTGAQKQISTRGQSLGIRMANPMGEVETGAMSNSS 312
303 SWNTIDEEERELTRRHEDGEPWMSFDFLRHYSRLICNLTPDTLTSOTYKKKLTQMD 362
313 ENNAVEBPALHQQLMVRMEDEFPWMSFDFLRHYSRLICNLTPDALQSRFRKNTLTYD 372
363 GNNRRGSTAGGCRNYPATFWNPPQKIRLDETDDDYDGRSGCSFVLTALMQKRRRR 419
373 GSWRRGSTAGGCRNYPATFWNPPQKIRLDETDDDYDGRSGCSFVLTALMQKRRRR 432
420 RKGEDMHTIGTCIYEVEPEELSGQTNHLSKNFPLTNRARESDPFINREVLNFKLPG 479
433 RRGKDMETIGTAVYEVPEELGQPLHKKRDFPLANRARESEOPINIREVSTRFLPG 492
480 GEYILVPSFTFEPNKGDFCIRVSEKADQYAVDEIEANTL-EEFDISEDIDDDGRRLLFA 538
493 GEYIVVPSFTFEPNKGDFCIRVSEKADQYAVDEIEANTL-EEFDISEDIDDDGRRLLFA 552
539 AQLAGEBAEISAFELQTIIRRLVLAQRDICKSDGSIETCKINVDMLDSDSGKLKKEFYI 598
553 ROLAGEBAEISAFELQTIIRRLVLAQRDICKSDGSIETCKINVDMLDSDSGKLKKEFYI 612
599 ILMTIKQKQKTYRELDVDRSGTMSYEMRKALBEAGFPMQOLQVIVARPADQIITDF 658
613 VLMMIRINVLSTIFRRFDLDRSGMSAYEMRKALBEAGFPMQOLQVIVARPADQIITDF 672
659 DNFRCLVRLLETLFKIFKQLDPEMTGTIELDLISWL 695
673 DNFRCLVRLLETLFKIFKQLDPEMTGTIELDLISWL 709

```

RESULT 6

```

ID 09N185 PRELIMINARY; PRT; 716 AA.
AC 09N185;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Micromolar calcium-dependent neutral protease large subunit.
GN CAPN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN 1;
RP SEQUENCE FROM N.A.
RA Smith T.P.L., Casae E., Rexroad C.E., Kappes S.M., Keele J.W.;
RT "Bovine CAPN1 maps to a region containing a QTL for meat tenderness.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF221129; AAF32364.1; -.

```



```

Db 493 PGEYVVVSTEEPNKGDVFLVRFSEKSGAGTQELDDQVQANLPDEQVLEESEIDENFKSL 552
Qy 538 FAOLAGEBAISAFELQTLIRVLAKRODIDKSDGSISITCKIMVMDLSDSGKLGKEP 597
Db 553 FROLAGEBEMESVSKELRTILNKILSKHDLRTGTSLSBSCRMVMDRDNKGKLGVEP 612
Qy 598 YLMTKIOKYOKIYREIVDVSRTGTMNSYEMRKALBEAGFKMPCQLHQVIVAFADQLII 657
Db 613 NILMRIRNYLITFRKFDLDSKSGMSAYEMRAIEAPFKLKKKILHIIITRYSEPLAV 672
Qy 658 DFDNVRCLVRLLETFKIFKQOLDPENTGTIELDLISWL 695
Db 673 DFDNVRCLVRLLETFRFEEKLDTDLIDGVVTFDLFKWL 710

```

RESULT 8

```

Q8C2J1 PRELIMINARY; PRT; 713 AA.
ID Q8C2J1;
AC Q8C2J1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Calpain 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ND; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK088547; BAC40416.1;
SQ SEQUENCE 713 AA; 82152 MW; 3E1FCB4D5802B864 CRC64;

```

Query Match 65.3%; Score 2426; DB 11; Length 713;

Best local Similarity 62.6%; Pred. No. 7.4e-176;

Matches 436; Conservative 126; Mismatches 130; Indels 4; Gaps 3;

```

Qy 3 GAAKLAQDRAEAELGSHERAIKYLNQDYALRNECEAGTLFQDPSPAPISALGKFE 62
Db 13 GVSAGVOCKKDKELGLGHENAIXKLGQDYETLRARCIQSGVLPQDEAFPPVSHSLGRKE 72
Qy 63 LGSYSKTRGMWRKRPTEICADPOFIIGATRTDICOAGLDCWLLAIAISLTNEBIIA 122
Db 73 LGSYSKTYGIKMKRPTEIMSNPQFIYDGAATRTDICOAGLDCWLLAIAISLTNEBIIH 132
Qy 123 RVPLNLSGFOENVAGIFHFQWQYGEWVYVVDRLPTKQGLLFPVHABESSEPSALLE 182
Db 133 RVSPGOSFOGVAGIFHFQWQYGEWVYVVDRLPTKQGLLFPVHABESSEPSALLE 192
Qy 183 KAVAKINGCYEALSGATTEGFEDFTGIAEWYELKKPPNLFKIIQALQKSLGCSI 242
Db 193 KAVAKINGSYALSGCCCTSEAFEDFTGIVTEWYDQKAPSDIYQIILKALRGSLGCSI 252
Qy 243 DITSAADSEAITFOKLVKGAHAYVTGABEVESNGSLQKILIRINPWEVETGRWINDCP 302
Db 253 NISDIRDEAITFKVLVGHAYSVTGAKOVITYQGQGVNLIMRNWNGEVEKGPMSDSY 312
Qy 303 SMNTIDPEERELTRRHEDGFWMSFSFLLHYSLKLCNTLPDTLSDTYKKMKLTAMD 362
Db 313 EWNVNDPVERQLRVKMDGFWMSFRDPIREFTKLEICNTLPALKSRITLRNNTTYFE 372
Qy 363 GNMRRGSTAGGCRNYPNTFMNPOYLIKLEBED--EEDDESGETFLVGLIOKRROR 420
Db 373 GTMRGSGTAGGCRNYPATFWNPQFKRLKEVDADDDYDNBSGGCSFLALAMQKRRER 432
Qy 421 KMGEDMTITGFIYEVPEELSQQTNILSKNPFILNRARERSDTFINLREVLNFKLPG 480

```

```

Db 433 RFGEDMTITGFAVYQVPRELAGQF-VHLKQDFLANASRAQSEHFILREVSNRIRLPQG 491
Qy 481 EYIVPSTFERNKDGDCIVFSEKADYQAVUDEIEANL-BEEDISEDDIDGVRLLA 539
Db 492 EYIVPSTFERNKDGDLLEFFSEKAGTQELDDQVQANLPDEKVSSEEDIDNFKTLFS 551
Qy 540 QLAGEDEAIAFAELQTLIRVLAKRODIDKSDGSISITCKIMVMDLSDSGKLGKEFYI 599
Db 552 KLAGDDHEISVKEIQTILNKILSKHDLRTGTSLSBSCRMVMDRDNKGKLGVEPNI 611
Qy 600 LMTKIOKYOKIYREIVDVSRTGTMNSYEMRKALBEAGFKMPCQLHQVIVAFADQLIIDF 659
Db 612 LMRIRNYLITFRKFDLDSKSGMSAYEMRAIEAPFKLKKKILHIIITRYSEPLAVDF 671
Qy 660 DNFVRCVRLLETFKIFKQOLDPENTGTIELDLISWL 695
Db 672 DNFVRCVRLLETFRFEEKLDTDLIDGVVTFDLFKWL 707

```

RESULT 9

```

Q918G2 PRELIMINARY; PRT; 704 AA.
ID Q918G2;
AC Q918G2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Calpain 1 (BC 3.4.22.17)
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee K.C., Evans C.W.;
RT "Zebrafish calpain 1 (camp1): Identification, expression and
RT phylogenetic implications."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF282675; AAF82808.1;
DR HSSP; P04574; IALV.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease C2.
DR InterPro; IPR00169; SHPoc acsite.
DR Pfam; PF00667; Calpain III; 1.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain III; 1.
DR SMART; SM00230; Cyapc; 1.
DR SMART; SM00054; Efp; 2.
DR PROSITE; PS00018; EF_HAND_2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR KEGG; K04472; CALPAIN.
SQ SEQUENCE 704 AA; 80090 MW; E5FBC08C18A79A8 CRC64;

```

Query Match 64.1%; Score 2381; DB 13; Length 704;

Best local Similarity 62.5%; Pred. No. 1.9e-172;

Matches 436; Conservative 105; Mismatches 153; Indels 4; Gaps 2;

```

Qy 3 GAAKLAQDRAEAELGSHERAIKYLNQDYALRNECEAGTLFQDPSPAPISALGKFE 62
Db 8 GMAARLSQMDRAGAGLQGNNAKPLGQDYETLRAGSQGSRRLFPEDMFLAASSSLGFNE 67
Qy 63 LGSYSKTRGMWRKRPTEICADPOFIIGATRTDICOAGLDCWLLAIAISLTNEBIIA 122
Db 68 LGSYSKTYGIVWRMRPREMCAKRPQFIYDGAATRTDICOAGLDCWLLAIAISLTNEBIIH 127
Qy 123 RVPLNLSGFOENVAGIFHFQWQYGEWVYVVDRLPTKQGLLFPVHABESSEPSALLE 182
Db 128 RVCRGTGQDFSRAGIHFHFQWQYGEWVYVVDRLPTKQGLLFPVHABESSEPSALLE 187
Qy 183 KAVAKINGCYEALSGATTEGFEDFTGIAEWYELKKPPNLFKIIQALQKSLGCSI 242

```



```

Db 188 KAAKANGCEALSGSGTCEGFEFTGVTGEMVELKAPADLPISIGRAIERSLGCS1 247
Cc 243 DITSADSEAITFOKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRANDNC 302
Cc 248 DITSKEDMEANTFKLVKGAAYSVTGAEEVYKGNMTKLVIRINPMGEVMTGANSNR 307
Cc 303 SWNTIDPEERERLTRRHEDGPFMWSFSDFLRHYSRLICNLTPDTLTSDTYKCKKLTQMD 362
Cc 308 EMDNDRSVGRGLONRSEDEPFMWSFSDFLREFLREICNLTDALQASQVKKWSTANYN 367
Cc 363 GNNRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRROR 422
Cc 368 GEMRRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRROR 424
Cc 423 GEMMTTGFVIEVEPEELSGOTNHLKSNFELTNRAERSDPTINLEVNRLFPREX 482
Cc 425 GOMETITGPAIYVEPEELSGOVHLKRDFTLHASSARSELFINLEVSRRFLPAGEX 484
Cc 483 ILVSTFEENKDDFCIRFSEKKADYQAVDEIEANL-EEFDISEDDIDDGVARLPAQL 541
Cc 485 IIVSTFEPEPKADFLVRFVSEKPNSEEMDKVMAIPEGRHDESDQIDAGFSLPRL 544
Cc 542 AGEDDAISAELOTLIRVLAKRODYSDFSIETCKINVDMLSDSGGLKEFTILM 601
Cc 545 AGADMEISVTELOTLIRVLAKRODYSDFSIETCKINVDMLSDSGGLKEFTILM 604
Cc 602 TKLOKOKIYREIDVRSCTGNSYEMKALEBAGFKPCQOLHOVYARFADOLIIDFN 661
Cc 605 EKIIRYLOTRDHDVDSGMSYEMKALEBAGFKPCQOLHOVYARFADOLIIDFN 664
Cc 662 FVRCVLVLETLFKIYKOLDPENTGTIELDLISWLCFSV 699
Cc 665 FVSCVLVLETLFKIYKOLDPENTGTIELDLISWLCFSV 702

RESULT 10
Q64698 PRELIMINARY; PRT; 703 AA.
AC 064698;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Calpain, large (Catalytic) subunit (EC 3.4.22.17) (Calcium-activated
DE neutral proteinase) (CAMP) (Stomach-specific calcium-activated neutral
DE protease large subunit) (NCCL2).
GN CUS4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=STOMACH;
RX MEDLINE=93374936; PubMed=7690035;
RA Sorimachi H., Ishiura S., Suzuki K.;
RT "A novel tissue-specific calpain species expressed predominantly in
RT the stomach comprises two alternative splicing products with and
RT without Ca(2+)-binding domain";
RL J. Biol. Chem. 268:19476-19482(1993).
CC - FUNCTION: CALPAINS ARE CALCIUM-ACTIVATED NON-LYSOSOMAL THIO-
CC PROTEASES.
CC - SUBUNIT: HETERODIMER OF A LARGE (CATALYTIC) AND A SMALL
CC (REGULATORY) SUBUNIT.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC.
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; NCCL2 (SHOWN HERE) AND NCCL2', ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE STOMACH. LOW IN
CC INTESTINE AND HEART.
CC - MISCELLANEOUS: IN RAT THERE SEEMS TO BE 2 TYPES OF CALPAIN-
CC UBQUITOUS FORMS, CALPAIN I (MICRO-MOLES CA++ REQUIRING) AND
CC CALPAIN II (MILI-MOLE CA++ REQUIRING), AND TISSUE SPECIFIC FORMS,
CC CALPAIN P94 AND NCCL2. THE SMALL UNIT IS COMMON TO ALL FORMS.

```

```

Cc - MISCELLANEOUS: THIS PROTEIN SEEMS TO BIND TWO MOLES OF CALCIUM.
Cc - SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
Cc - SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2; ALSO KNOWN AS THE
Cc CALPAIN FAMILY OF THIOL PROTEASES.
Cc EMBL: D14478; BA03369.1; -
Cc EMBL: D14479; BA03370.1; -
Cc EMBL: D14480; BA03371.1; -
Cc HSSP: P04574; IALV.
Cc DR MEROPS: C02.007; -.
Cc DR InterPro: IPR002048; EF-hand.
Cc DR InterPro: IPR001300; Protease_C2.
Cc DR InterPro: IPR000169; SHpro_csite.
Cc DR Pfam: PF01067; Calpain_III; 1.
Cc DR Pfam: PF00036; efnhand; 4.
Cc DR Pfam: PF00648; Peptidase_C2; 1.
Cc DR PRINTS: PR00704; CALPAIN.
Cc DR SMART: SM00720; calpain_III; 1.
Cc DR SMART: SM00230; CyBpc; 1.
Cc DR SMART: SM00054; Efn; 2.
Cc DR PROSITE: PS00018; EF_HAND; 1.
Cc DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
Cc DR Hydrolase: Thiol protease; Calcium-binding; Multigene family;
Cc Alternative splicing.
Cc FT DOMAIN 1 73
Cc FT 74 332 I, REGULATION OF PROTEASE ACTIVITY.
Cc FT DOMAIN 333 532 II, THIOL PROTEASE.
Cc FT DOMAIN 533 703 IV, CALCIUM-BINDING.
Cc FT ACT_SITE 105 105 BY SIMILARITY.
Cc FT ACT_SITE 262 262 BY SIMILARITY.
Cc FT ACT_SITE 286 286 BY SIMILARITY.
Cc FT CA_BIND 588 599 SITE II (PROBABLE).
Cc FT CA_BIND 618 629 SITE II (PROBABLE).
Cc FT DOMAIN 653 664 ANCESTRAL CALCIUM SITE III (POTENTIAL).
Cc FT DOMAIN 683 694 ANCESTRAL CALCIUM SITE IV (POTENTIAL).
Cc FT VARSPIC 103 103 G -> V (IN ISOFORM NCCL2').
Cc FT VARSPIC 380 380 T -> S (IN ISOFORM NCCL2').
Cc FT VARSPIC 381 703 MISSING (IN ISOFORM NCCL2').
Cc SQ SEQUENCE 703 AA; 79555 MW; C0688B055FC0DEC CRC64;

Query Match 63.5%; Score 2356.5; DB 11; Length 703;
Best Local Similarity 60.9%; Pred. No. 1.4e-170; Indels 7; Gaps 3;
Matches 427; Conservative 127; Mismatches 140;

Cc 1 MAGIAKLAKDRRAAAGLSGHERAIVKYNQDYALNNECLAGTLRQDPSFPAIPALGF 60
Cc 1 MALAAGVSKQRAVAGLSGNSQNAVYLGDEFTLAKQCLNSVLRKDPFRPCBALGV 60
Cc 121 LARVPLNGSFQNNYAGIFHFQFQWQGEWVVDRLPTKDELLFVHABSEFMSAL 180
Cc 121 LYRVLPRDPSFQDYAGIFHFQFQWQGEWVVDRLPTKDELLFVHABSEFMSAL 180
Cc 181 LERAYAKINGCYALSGATTEGFEFTGIAEWEYELKRPPIFKTIQALQKSLG 240
Cc 181 LERAYAKINGCYALSGATTEGFEFTGIAEWEYELKRPPIFKTIQALQKSLG 240
Cc 241 SIDITSADSEAITFOKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRAND 300
Cc 241 SIDITSADSEAITFOKLVKGAAYSVTGAEEVNSGLOKLRIRNPMGEVMTGRAND 300
Cc 301 CPBWNTIDPEERERLTRRHEDGPFMWSFSDFLRHYSRLICNLTPDTLTSDTYKCKKLT 360
Cc 301 CPBWNTIDPEERERLTRRHEDGPFMWSFSDFLRHYSRLICNLTPDTLTSDTYKCKKLT 360
Cc 361 MDGNRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRR 418
Cc 361 MDGNRSGTAGCCNRYPTFMNPOYLKLEEBDEDEDESGCTFLVGLIQKRRR 418
Cc 419 QRMKEDMTTGFVIEVEPEELSGOTNHLKSNFELTNRAERSDPTINLEVNRLPKLP 478
Cc 419 QRMKEDMTTGFVIEVEPEELSGOTNHLKSNFELTNRAERSDPTINLEVNRLPKLP 478

```

Db 421 QKRIGQMLSIGAVYQPKELSHSTDAHLGRDFELGRQPSSTSYNNLEVSRYRLP 480
Qy 479 PGEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLEEF--FDISEDDIDGVR 535
Db 481 PGOYLVPSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLEEF--DEIVR 538
Qy 536 RLFAQLAGEADAISAFELQTLIRRLARLAKODIKSDGFSIETCKIMVMDLSDSGKLGK 595
Db 539 SLPEFVQKDEISANQLKRYLNEVLKRTDMKFDGFINTCREMIISLSDSGTSLGPM 598
Qy 596 EFTYLMKIRYQKRYREIDVDRSGTMSYEMRKALAEAGKRCQHLQHVAVFADQL 655
Db 599 EFTYLMKIRYQKRYREIDVDRSGTMSYEMRKALAEAGKRCQHLQHVAVFADQL 658
Qy 656 IIDPNEVRCIVRLETLFKIFKQDPENTGTIELDLISMLC 696
Db 659 GVDINGFVACMIRLETLFKIFRLDKONGIYQSLAEMLC 699

RESULT 11

Q8BPV9 PRELIMINARY; PRT; 530 AA.
AC Q8BPV9; 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Calpain 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MBLINB=22354683; Pubmed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK052141; BAC34855.1; -
DR SEQUENCE 530 AA; 5929 MW; 7BD6DCDCC1482BC92 CRC64;
Query Match 62.9%; Score 2335; DB 11; Length 530;
Best Local Similarity 71.0%; Pred. No. 4e-169;
Matches 451; Conservative 24; Mismatches 28; Indels 132; Gaps 2;

Qy 1 MAGIAAKIADKREAEGLGSHERAIKYINODYEALRNECLAGTLFQDPSPPAISALGF 60
Db 1 MAGIAAKIADKREAEGLGSHERAIKYINODYEALRNECLAGTLFQDPSPPAISALGF 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALDPCWLLAASLTINEI 120
Db 61 KELGPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALDPCWLLAASLTINEI 120
Qy 121 LARVVPINOSPOENVAGIFHFQWQYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Db 121 LARVVPINOSPOENVAGIFHFQWQYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Qy 121 LARVVPINOSPOENVAGIFHFQWQYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Db 121 LARVVPINOSPOENVAGIFHFQWQYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Qy 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSGSLGC 240
Db 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSGSLGC 240
Qy 241 SIDITSAADSAIFQKLVKHAHSVYGAEEVESNGSIQKILIRNPMGEVEMTKRWN 300
Db 241 SIDITSAADSAIFQKLVKHAHSVYGAEEVESNGSIQKILIRNPMGEVEMTKRWN 300
Qy 301 CPSNNTIDPEERERLTRHDEGEFPMFSDFLRHYSRLCNLPDITLSDTYKKMKLTK 360
Db 301 CPSNNTIDPEERERLTRHDEGEFPMFSDFLRHYSRLCNLPDITLSDTYKKMKLTK 360
Qy 361 MDGNMRGSGTAGGCRNPNTPMNPQVYLKLEEDDEDEDESGCTPLVGLIQRKRROR 420

Db 359 ----- 358
Qy 421 KMGEWMHTIGFIYEVEBELSGQTNHLSKNFELTNBARERSDFTINLREVLRFLKLP 480
Db 359 ----- 358
Qy 481 EYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLEEF--FDISEDDIDGVRRLAQ 540
Db 359 -----TFPHKDGDFCIRVFSEKKADYQAVDEIEANLEEF--IDANEEDIDGFRRLVQ 411
Qy 541 LAGEADAISAFELQTLIRRLARLAKODIKSDGFSIETCKIMVMDLSDSGKLGKXERYIL 600
Db 412 LAGEADAISAFELQTLIRRLARLAKODIKSDGFSIETCKIMVMDLSDSGKLGKXERYIL 471
Qy 601 WTKIQYQKRYREIDVDRSGTMSYEMRKALAEAG 635
Db 472 WTKIQYQKRYREIDVDRSGTMSYEMRKALAEAG 503

RESULT 12

Q91VA3 PRELIMINARY; PRT; 703 AA.
AC Q91VA3; 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Stomach specific calpain nCL-2 (Stomach-specific calpain).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hata S., Sorimachi H., Lee H., Kawahara H., Maeda T., Suzuki K.;
RT "Both conserved and unique gene structure of stomach-specific calpains
RT revealed a process of calpain gene evolution."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Hata S., Nishi K., Kawamoto T., Lee H., Kawahara H., Maeda T.,
RA Shintani Y., Sorimachi H., Suzuki K.;
RT "Both the conserved and unique gene structure of stomach-specific
RT calpains reveal processes of calpain gene evolution."
RL J. Mol. Evol. 0:0-0(2001).
DR EMBL; AB050202; BAB70480.1; -
DR EMBL; AB061518; BAB55000.1; -
DR MEROPS; C02.007; -
DR MGD; MGI:2181366; Capn8.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Protease C2.
DR InterPro; IPR001659; SHprot_acsite.
DR Pfam; PF01067; Calpain_III; 1.
DR Pfam; PF00036; ehand; 3.
DR Pfam; PF00648; peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; Cyapc; 1.
DR SMART; SM00054; Efn; 3.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; 1.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; 1.
DR SEQUENCE 703 AA; 50F2P9CCF2192PB6 CRC64;

Query Match 62.6%; Score 2322.5; DB 11; Length 703;
Best Local Similarity 60.2%; Pred. No. 5.5e-168;
Matches 421; Conservative 131; Mismatches 144; Indels 3; Gaps 2;

Qy 1 MAGIAAKIADKREAEGLGSHERAIKYINODYEALRNECLAGTLFQDPSPPAISALGF 60
Db 1 MAGIAAKIADKREAEGLGSHERAIKYINODYEALRNECLAGTLFQDPSPPAISALGF 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALDPCWLLAASLTINEI 120
Db 61 KELGPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALDPCWLLAASLTINEI 120
Qy 121 LARVVPINOSPOENVAGIFHFQWQYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Db 121 LARVVPINOSPOENVAGIFHFQWQYGEWVWVDDRLPTQDGLLEFVHSAEGSEFMSAL 180
Qy 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSGSLGC 240
Db 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSGSLGC 240
Qy 241 SIDITSAADSAIFQKLVKHAHSVYGAEEVESNGSIQKILIRNPMGEVEMTKRWN 300
Db 241 SIDITSAADSAIFQKLVKHAHSVYGAEEVESNGSIQKILIRNPMGEVEMTKRWN 300
Qy 301 CPSNNTIDPEERERLTRHDEGEFPMFSDFLRHYSRLCNLPDITLSDTYKKMKLTK 360
Db 301 CPSNNTIDPEERERLTRHDEGEFPMFSDFLRHYSRLCNLPDITLSDTYKKMKLTK 360
Qy 361 MDGNMRGSGTAGGCRNPNTPMNPQVYLKLEEDDEDEDESGCTPLVGLIQRKRROR 420
Db 361 MDGNMRGSGTAGGCRNPNTPMNPQVYLKLEEDDEDEDESGCTPLVGLIQRKRROR 420

QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLKPPNLFKTIQKALQKSLG 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLKPPNLFKTIQKALQKSLG 240
 QY 241 SIDTSAADSEAITFOKLVKHAYSVTGAEEVESNGSLQKLIBIRNPMGEVETGWRWNN 300
 DB 241 SIDTSAADSEAITFOKLVKHAYSVTGAEEVESNGSLQKLIBIRNPMGEVETGWRWNN 300
 QY 301 CPSEMTIDPEERERLTRRHDEGEFMSFSDPLRHYSRLICNTLPTDITSDTYKKMLTK 360
 DB 301 CPSEMTIDPEERERLTRRHDEGEFMSFSDPLRHYSRLICNTLPTDITSDTYKKMLTK 360
 QY 361 MDGWRGSTRAGGCRNYPNTFMNPOYLKLEEDDEDEB--ESGCTFLVGLIQKRRR 418
 DB 361 MDGWRGSTRAGGCRNYPNTFMNPOYLKLEEDDEDEB--ESGCTFLVGLIQKRRR 418
 QY 419 QKKGEDMTIGFIYVEPEELSGQTNILSKNFFLTNRARESDFTLREVLNRPFLP 478
 DB 419 QKKGEDMTIGFIYVEPEELSGQTNILSKNFFLTNRARESDFTLREVLNRPFLP 478
 QY 421 QRRGQMLSTIGYAVYQIKPELENHTDEHIGRDFQGRQSTGCTTWNLRVSSRVQLP 480
 DB 421 QRRGQMLSTIGYAVYQIKPELENHTDEHIGRDFQGRQSTGCTTWNLRVSSRVQLP 480
 QY 479 PGEYILVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 537
 DB 479 PGEYILVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 537
 QY 481 PGQVLVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 540
 DB 481 PGQVLVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 540
 QY 538 PAQAGADAEISAFELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 597
 DB 538 PAQAGADAEISAFELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 597
 QY 541 SEEDADSEISAFELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 600
 DB 541 SEEDADSEISAFELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 600
 QY 598 YILWTIKQYKIVREIDVRSQTNISYEMKALEAGFKMPQOLHOVIVARFADDOILI 657
 DB 598 YILWTIKQYKIVREIDVRSQTNISYEMKALEAGFKMPQOLHOVIVARFADDOILI 657
 QY 601 KTLMLKICXLEIYQENHDSBAGIIDAHEKRTALKKAGFTLANNVOQTITATRIACSKLGV 660
 DB 601 KTLMLKICXLEIYQENHDSBAGIIDAHEKRTALKKAGFTLANNVOQTITATRIACSKLGV 660
 QY 658 FDNFVRCVLVLETLFKIQOLDPENTGTIELDLISWLC 696
 DB 658 FDNFVRCVLVLETLFKIQOLDPENTGTIELDLISWLC 696
 QY 661 FDFGVACMIRLETLFKIQOLDPENTGTIELDLISWLC 699
 DB 661 FDFGVACMIRLETLFKIQOLDPENTGTIELDLISWLC 699

RESULT 13
 Q918T0 PRELIMINARY; PRT; 702 AA.
 ID Q918T0; PRELIMINARY; PRT; 702 AA.
 AC Q918T0; PRELIMINARY; PRT; 702 AA.
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Calpain.
 GN CL-2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCBITaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT Cao Y., Zhao H., Grunz H.:
 RT "A novel Xenopus gene homologous to rat calpain.",
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF212199; AF63194.2; -.
 DR HSSP; P04574; 1ALV.
 DR MEROPS; C02.007; -.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR000169; SHprot_acstle.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN_III; 1.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyapc; 1.

DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR SEQUENCE 702 AA; 7946 MW; 6D66A3B1D276DC51 CRC64;
 Query Match 62.2%; Score 2309; DB 13; Length 702;
 Best Local Similarity 60.7%; Pred. No. 5.8e-167;
 Matches 426; Conservative 119; Mismatches 155; Indels 2; Gaps 1;
 QY 1 MAGIAKLADEAEGISHERAIKYNODYALNRECLEAGTLFODPEFPPIPSALGF 60
 DB 1 MMSAANIAADRLADGCGIKRNPETKLDDEFKLAQCLASGALYDEFPACPSALGY 60
 QY 61 KELGPYSSKTRGMRKRPTEICADPOITIGATRTDICGALDCWLLAIALSLTNEE 120
 DB 61 NELRPGSYKTSGYIWKRPTEICPNPQIYVDAATRGDIRQALDCWLLAIALSLTNEE 120
 QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVVDRLPTKQGLLFVHSAEGSEFMSAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLKPPNLFKTIQKALQKSLG 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEDEFTGIAEMYLKPPNLFKTIQKALQKSLG 240
 QY 241 SIDTSAADSEAITFOKLVKHAYSVTGAEEVESNGSLQKLIBIRNPMGEVETGWRWNN 300
 DB 241 SIDTSAADSEAITFOKLVKHAYSVTGAEEVESNGSLQKLIBIRNPMGEVETGWRWNN 300
 QY 241 SIDTNAVDEATLSTRKLVKHAYSVTGAEEVLVYRGRQELIRVNPGEVETGWRWNN 300
 DB 241 SIDTNAVDEATLSTRKLVKHAYSVTGAEEVLVYRGRQELIRVNPGEVETGWRWNN 300
 QY 301 CPSEMTIDPEERERLTRRHDEGEFMSFSDPLRHYSRLICNTLPTDITSDTYKKMLTK 360
 DB 301 CPSEMTIDPEERERLTRRHDEGEFMSFSDPLRHYSRLICNTLPTDITSDTYKKMLTK 360
 QY 301 APEMNIVDPYKAVLDRKSEDEGFMAFSDPLREYSRLICNTLPTDITSDTYKKMLTK 360
 DB 301 APEMNIVDPYKAVLDRKSEDEGFMAFSDPLREYSRLICNTLPTDITSDTYKKMLTK 360
 QY 361 MDGWRGSTRAGGCRNYPNTFMNPOYLKLEEDDEDEB--DSEGCTFLVGLIQKRRR 418
 DB 361 MDGWRGSTRAGGCRNYPNTFMNPOYLKLEEDDEDEB--DSEGCTFLVGLIQKRRR 418
 QY 361 YTSMAKSGSTRAGGCRNYPNTFMNPOYLKLEEDDEDEB--DSEGCTFLVGLIQKRRR 420
 DB 361 YTSMAKSGSTRAGGCRNYPNTFMNPOYLKLEEDDEDEB--DSEGCTFLVGLIQKRRR 420
 QY 419 QKKGEDMTIGFIYVEPEELSGQTNILSKNFFLTNRARESDFTLREVLNRPFLP 478
 DB 419 QKKGEDMTIGFIYVEPEELSGQTNILSKNFFLTNRARESDFTLREVLNRPFLP 478
 QY 421 KKKMGEDLSTIGSLFKIPQOLDHDAHIGRDFQGRQSTGCTTWNLRVSSRVQLP 480
 DB 421 KKKMGEDLSTIGSLFKIPQOLDHDAHIGRDFQGRQSTGCTTWNLRVSSRVQLP 480
 QY 479 PGEYILVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 538
 DB 479 PGEYILVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 538
 QY 481 VGDYLVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 540
 DB 481 VGDYLVSTFEENKQDFCIRVSEKKADYQAVDEIEANLEB-PDISBDIDGVRRL 540
 QY 539 AOLAGDAEISAFELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 598
 DB 539 AOLAGDAEISAFELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 598
 QY 541 DKLAGDEEVDARELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 600
 DB 541 DKLAGDEEVDARELQTLIRVLAQKODIKDSFSETCKIMVMDSDSGKLGKEF 600
 QY 599 ILWTIKQYKIVREIDVRSQTNISYEMKALEAGFKMPQOLHOVIVARFADDOILI 658
 DB 599 ILWTIKQYKIVREIDVRSQTNISYEMKALEAGFKMPQOLHOVIVARFADDOILI 658
 QY 601 ILWMKIQKVAITLKKADSDSGIMDSHEKRTALKKAGFTLANNVOQTITATRIACSKLGV 660
 DB 601 ILWMKIQKVAITLKKADSDSGIMDSHEKRTALKKAGFTLANNVOQTITATRIACSKLGV 660
 QY 659 FDNFVRCVLVLETLFKIQOLDPENTGTIELDLISWLC 700
 DB 659 FDNFVRCVLVLETLFKIQOLDPENTGTIELDLISWLC 700
 QY 661 FDFGVACMIRLETLFKIQOLDPENTGTIELDLISWLC 702
 DB 661 FDFGVACMIRLETLFKIQOLDPENTGTIELDLISWLC 702

RESULT 14
 Q9NOM7 PRELIMINARY; PRT; 650 AA.
 ID Q9NOM7; PRELIMINARY; PRT; 650 AA.
 AC Q9NOM7; PRELIMINARY; PRT; 650 AA.
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Microtalar calcium-activated neutral protease I isoform B.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NC NCBITaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT MEDLINE=21087425; PubMed=11219468;
 RX Smith T.P.L., Simmen F.A., Zhao G., Vallet J.L.;

"Nucleotide sequences of two isoforms of porcine micromolar calcium-RT activated neutral protease 1 cDNA."
 RT J. Anim. Sci. 79:552-553(2001).
 DR EMBL; AF263609; AAF73443.1; -
 DR HSPSP; P04574; IALV.
 DR MEROPS; C02.001; -
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR000169; SHprot_acidic.
 DR Pfam; PF01067; Calpain_III_1.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyapc; 1.
 DR SMART; SM00054; Efh; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOI_PROTEASE_CYS; 1.
 DR KMW
 SEQUENCE 650 AA; 74033 MW; 2E9E9F569D90FD7 CRC64;

Query Match 61.4%; Score 2279.5; DB 6; Length 650;
 Best Local Similarity 64.5%; Pred. No. 9.1e-165;
 Matches 411; Conservative 107; Mismatches 114; Indels 5; Gaps 3;

QY 3 GIAAKLAKDRAEAGLSHERAIKYNQYALNNECLEAGTLQDPSPFPAIPSLGPKF 62
 13 GVSAGVQQLRAKEIGLGHENAIKYLQGVFQLRHCHQSSLPFDEAFPPVQSLGKE 72
 QY 63 LGPYSSKTRGMKRPTEICADPOFICGATRTDICQALGDCWLLAIAISLTNEIILA 122
 73 LGPNSSKTYGVKMKRPTELFSNPQFIVDGAATRTDICQALGDCWLLAIAISLTNDLILH 132
 QY 123 RVVLNQSFOENVAGIFHFQFQMGWEVVDRLPTKDELLFVHSAGSEFMSALLE 182
 133 RVPHGQSFQNGVAGIFHFQFQMGWEVVDRLPTKDELLFVHSAGSEFMSALLE 192
 QY 183 KAVAKINGCYALSGGATTEGFEDFTGIAEMYLKRPNNLFXIIQALQKSLGCSI 242
 193 KAVAKINGCYALSGGATTEGFEDFTGIAEMYLKRPNNLFXIIQALQKSLGCSI 252
 QY 243 DITSAADSEAITPQKLVGHAIVSTGAEEVSNGLQKLIIRINPWGEVENTGRANDNC 302
 253 DISSVLDMEAIVTFFKLVGHAIVSTGAEEVSNGLQKLIIRINPWGEVENTGRANDNC 312
 QY 303 SMTNIDPERERLTPRHEDGFEWMSFSDFLRHSLEICNLTPDTLTSDTYKKMLTKMD 362
 313 EWNQVDPYQDQLRVRMEDGFEWMSFSDFLRHSLEICNLTPDTLTSDTYKKMLTKMD 372
 QY 363 GNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEED---GESGCTFLVGLIQHRRQ 419
 373 GNMRRGSTAGCCRNYPNTFMNPOYLKLEEDDEED---GESGCTFLVGLIQHRRQ 432
 QY 420 RKMGEWHITIGFIYVEPEELSGQTNHLSKNFELTNRAERSDTFINLREVLNFKLP 479
 433 RRFQDMETIGFIYVEPEELSGQTNHLSKNFELTNRAERSDTFINLREVLNFKLP 491
 QY 480 GEYLIVSTFEPNKGDCIFVSEKADYQAVDEIEANL-EEEDISEDIDDOVRLLF 538
 492 GEYLIVSTFEPNKGDCIFVSEKADYQAVDEIEANL-EEEDISEDIDDOVRLLF 551
 QY 539 AQLAGEDAIEISAFELQTLIRVLAKRQDIKSDGFIETCKIMVMDLSDSGSLGLKEFY 598
 552 RQLAGEDEIEISAFELQTLIRVLAKRQDIKSDGFIETCKIMVMDLSDSGSLGLKEFY 611
 QY 599 ILMTKIQYQKIYREIVDQSGTMSYEMRKALEAG 635
 612 ILMTKIQYQKIYREIVDQSGTMSYEMRKALEAG 648
 Db

RESULT 15
 Q8R486 PRELIMINARY; PRT; 421 AA.

AC Q8R486;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE M-calpain 80 kDa large subunit (Fragment).
 GN CAPN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RA Elce J.S., Arthur J.S.C., Croall D.E., Dutt P.;
 RT "Mouse m-calpain, partial genomic sequence."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF97625; AAM19226.1; -
 DR InterPro; IPR001300; Protease_C2.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyapc; 1.
 FT NON_TER 1
 FT 421 1
 FT 421 1
 SQ SEQUENCE 421 AA; 48334 MW; 27E450B7B95E633 CRC64;

Query Match 57.7%; Score 2142; DB 11; Length 421;
 Best Local Similarity 93.3%; Pred. No. 1.4e-154;
 Matches 393; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

QY 143 FMQGEWEVVDRLPTKDELLFVHSAGSEFMSALLEKAVAKINGCYEALSGATTE 202
 1 FMQGEWEVVDRLPTKDELLFVHSAGSEFMSALLEKAVAKINGCYEALSGATTE 60
 QY 203 GFEDFTGIAEMYLKRPNNLFXIIQALQKSLGCSIDITSAADSEAITPQKLVKH 262
 61 GFEDFTGIAEMYLKRPNNLFXIIQALQKSLGCSIDITSAADSEAITPQKLVKH 120
 QY 263 AYSTGAEEVSNGLQKLIIRINPWGEVENTGRANDNCPSMTNIDPERERLTPRHEDG 322
 121 AYSTGAEEVSNGLQKLIIRINPWGEVENTGRANDNCPSMTNIDPERERLTPRHEDG 180
 QY 323 EFWMSFSDFLRHSLEICNLTPDTLTSDTYKKMLTKMDGNRRGSTAGCCRNYPNTFW 382
 181 EFWMSFSDFLRHSLEICNLTPDTLTSDTYKKMLTKMDGNRRGSTAGCCRNYPNTFW 240
 QY 383 MNPQYLLKLEEDDEEDGSGCTFLVGLIQHRRQKMGEDHTIGFIYVEPEELSG 442
 241 MNPQYLLKLEEDDEEDGSGCTFLVGLIQHRRQKMGEDHTIGFIYVEPEELSG 300
 QY 443 QTNHLSKNFELTNRAERSDTFINLREVLNFKLPGEYLIVSTFEPNKGDCIFCIYRV 502
 301 QTNHLSKNFELTNRAERSDTFINLREVLNFKLPGEYLIVSTFEPNKGDCIFCIYRV 360
 QY 503 SEKKADYQAVDEIEANL-EEEDISEDIDDOVRLLFQLAGDAIEISAFELQTLIRVLA 562
 361 SEKKADYQAVDEIEANL-EEEDISEDIDDOVRLLFQLAGDAIEISAFELQTLIRVLA 420
 QY 563 K 563
 Db 421 K 421

Search completed: July 24, 2003, 12:56:57
 Job time : 89.5929 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:28:44 ; Search time 19.3504 Seconds

(without alignments)
1701.191 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713

Sequence: 1 MAGIAAKAKDREAEGLGS.....PENTGTIELDLISWLCPSVL 700

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3704	99.8	700	1	CAN2_HUMAN
2	3522	94.9	700	1	CAN2_MOUSE
3	3518	94.7	700	1	CAN2_RAT
4	3098	83.4	700	1	CAN2_CHICK
5	2579.5	69.5	705	1	CAN2_CHICK
6	2458.5	66.2	714	1	CAN1_HUMAN
7	2444.5	65.8	714	1	CAN1_MOUSE
8	2435	65.6	713	1	CAN1_MOUSE
9	2427	65.4	713	1	CAN1_RAT
10	2153	58.0	422	1	CAN2_RABIT
11	2014	54.2	702	1	CAN2_HUMAN
12	1897	51.1	810	1	CAN3_CHICK
13	1877.5	50.6	821	1	CAN3_MOUSE
14	1869.5	50.4	821	1	CAN3_RAT
15	1862.5	50.2	821	1	CAN2_HUMAN
16	1619	43.6	324	1	CAN2_PIG
17	1594.5	43.7	828	1	CAN1_MOUSE
18	1253	32.9	783	1	YKR2_CAEEL
19	1202.5	32.4	758	1	CAN1_MOUSE
20	965.5	25.8	640	1	CAN2_MOUSE
21	957	25.0	207	1	CAN2_BOVIN
22	932.5	25.1	640	1	CAN5_RAT
23	931	25.1	640	1	CAN5_HUMAN
24	925.5	24.9	302	1	CAN1_RABIT
25	790	21.3	374	1	CAN3_PIG
26	788	21.2	641	1	CAN6_HUMAN
27	780	21.0	641	1	CAN6_MOUSE
28	772	20.8	641	1	CAN6_RAT
29	634	17.1	666	1	CANA_RAT
30	627.5	16.9	666	1	CANA_MOUSE
31	609	16.4	653	1	CANA_MACFA
32	604	16.3	672	1	CANA_HUMAN
33	458	12.3	263	1	CAN5_BOVIN

34	458	12.3	266	1	CAN5_PIG	P04574 sus scrofa
35	454	12.2	266	1	CAN5_RABIT	P06813 oryctolagus
36	453	12.2	269	1	CAN5_MOUSE	P08456 mus musculus
37	450	12.1	268	1	CAN5_HUMAN	P04632 homo sapien
38	445	12.0	266	1	CAN5_RAT	P04537 ratius norv
39	426	11.5	1597	1	SOL_DROME	P27388 drosophila
40	240	6.5	198	1	SORC_HUMAN	P30626 homo sapien
41	239	6.4	217	1	GRAN_HUMAN	P28676 homo sapien
42	236	6.4	198	1	SORC_CRIO	P05044 cricetus
43	195	5.3	116	1	CAN3_BOVIN	P51186 bos taurus
44	178	4.8	171	1	SORC_SCHUA	P04743 schistosoma
45	170	4.6	191	1	PCDE_MOUSE	P12815 mus musculus

ALIGNMENTS

RESULT 1

ID	CAN2_HUMAN	STANDARD;	PRT;	700 AA.
AC	P17655; Q16738; Q8WU26; Q9HBB1;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Calpain 2, large [catalytic] subunit precursor (BC 3.4.22.17)			
DE	(Calcium-activated neutral proteinase) (CANP) (M-type) (M-calpain)			
DE	(Molluscal-calpain) (Calpain large polypeptide I2).			
GN	CANP2 OR CANP12			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8916474; PubMed=2852952;			
RA	Imajo S., Aoki K., Ono S., Emori Y., Kawasaki H., Sugihara H.,			
RA	Suzuki K.;			
RT	"Molecular cloning of the cDNA for the large subunit of the			
RT	high-Ca2+-requiring form of human Ca2+-activated neutral protease.";			
RL	Biochemistry 27:8122-8126(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Datchenko L., Marushta K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,			
RA	Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Murthy D.M., Sodegryn E.J., Lu X., Gibbs R.A.,			
RA	Felby J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,			
RA	Schuerch A., Schein J.E., Jones S.J.M., Maira M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RL	[4]			

Query Match 99.8%; Score 3704; DB 1; Length 700;
 Best Local Similarity 99.7%; Pred. No. 1,1e-22;
 Matches 699; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGIAKLADREAREAGISHERAIXKYNODYALNRECEAGTLPODPSFPALPSALGF 60
 DB 1 MAGIAKLADREAREAGISHERAIXKYNODYALNRECEAGTLPODPSFPALPSALGF 60
 QY 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWMLAIASTLTNEE 120
 DB 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCWMLAIASTLTNEE 120
 QY 121 LARAVPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 180
 DB 121 LARAVPLNOSFOENYAGIFHFQWQYGEWVEVVDRLPTKDGELLFVHSAEGSEFWAL 180
 QY 181 LERAVKINCCYALSGGATTEGEBEDTGGIAEYELKPPMLFKLIQALQKSLG 240
 DB 181 LERAVKINCCYALSGGATTEGEBEDTGGIAEYELKPPMLFKLIQALQKSLG 240
 QY 241 SIDITSAADSEALTFOKLVGHAYSVTGAEEVNSGSLQKILIRNPMGEVMTGRMND 300
 DB 241 SIDITSAADSEALTFOKLVGHAYSVTGAEEVNSGSLQKILIRNPMGEVMTGRMND 300
 QY 301 CPSEWNTIDPERERLTRRHEDGFWMSSDFLRHYSRLKICNLTPTDITSDIYKKWLT 360
 DB 301 CPSEWNTIDPERERLTRRHEDGFWMSSDFLRHYSRLKICNLTPTDITSDIYKKWLT 360
 QY 361 MDONMRGSTAGCCRYNPTFMNPPQVLILEEDEDGESEGCFTVLGLOKRRROR 420
 DB 361 MDONMRGSTAGCCRYNPTFMNPPQVLILEEDEDGESEGCFTVLGLOKRRROR 420
 QY 421 KMGEDMHTTGFQYVEYELSGQTNHLSKNFPLNPARSRSTFINLREVLNRFKLP 480
 DB 421 KMGEDMHTTGFQYVEYELSGQTNHLSKNFPLNPARSRSTFINLREVLNRFKLP 480
 QY 481 EYLLVSTFEFPNKGDFCIRVFSEKADYQAVDEIEANLEEDDISDDIDGVRLLFAQ 540
 DB 481 EYLLVSTFEFPNKGDFCIRVFSEKADYQAVDEIEANLEEDDISDDIDGVRLLFAQ 540
 QY 541 LAEDEDAISAFELQTLIRRLAKRODKSGFSIFETKINWMLDSGSGKLGKERYIL 600
 DB 541 LAEDEDAISAFELQTLIRRLAKRODKSGFSIFETKINWMLDSGSGKLGKERYIL 600
 QY 601 WTKIOYKOKIYREIDVDRSGTMSYEMRKALBEAGFMPQCLHOVIVARFADQLIIDP 660
 DB 601 WTKIOYKOKIYREIDVDRSGTMSYEMRKALBEAGFMPQCLHOVIVARFADQLIIDP 660
 QY 661 NFVRCIVRLLETLFKIFKQIDPENTGTELDLISMLCPSVL 700
 DB 661 NFVRCIVRLLETLFKIFKQIDPENTGTELDLISMLCPSVL 700

RA Dear T.N., Matena K., Vingron M., Boehm T.;
 RT "A new subfamily of vertebrate calpains lacking a calmodulin-like
 RL domain: implications for calpain regulation and evolution.";
 RN Genomics 45:175-184 (1997).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Oaki Y.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=CNS;
 RA Glaes J.D., Naeh N.R., Dry I., Culver D., Wesselingh S.;
 RT "Cloning of m-calpain from mouse nervous system";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyzes limited proteolysis of substrates involved in
 CC cytoskeletal remodeling and signal transduction (by similarity).
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the p2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
 CC of calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding.
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL; Y10139; CA471227.1; -;
 DR EMBL; D38117; BAA32964.1; -;
 DR EMBL; AF015038; AAB94029.1; -;
 DR HSSP; Q07009; IDFO.
 DR MEROPS; C02.002; -;
 DR MGD; MGI:88264; Capn2.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR000169; SHPoc_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; efhand; 2.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cypec; 1.
 DR SMART; SM00054; Eph; 2.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KW Hydrolyase; Thiol protease; Calcium-binding; Repeat; Multisene family.
 FT CHAIN 1 19
 FT PROPEP 1 19
 FT DOMAIN 20 700
 FT DOMAIN 20 355
 FT DOMAIN 356 514
 FT DOMAIN 515 528
 FT DOMAIN 530 700
 FT CA_BIND 541 552
 FT CA_BIND 585 596
 FT CA_BIND 615 626
 FT DOMAIN 650 661
 FT DOMAIN 680 691
 FT ACT_SITE 105 109
 FT ACT_SITE 262 262
 FT ACT_SITE 286 286
 FT ACT_SITE 286 286
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.

FT	CONFLICT	194	194	A -> T (IN REF. 1).	
FT	CONFLICT	212	212	A -> G (IN REF. 1).	
FT	CONFLICT	402	402	E -> G (IN REF. 1).	
SQ	SEQUENCE	700 AA;	79871 MM;	682146B290968316 CRC64;	
Query Match					
Best Local Similarity		94.9%; Score 3522; DB 1; Length 700;			
Matches		655;	Conservative	24;	Mismatches 21; Indels 0; Gaps 0
QY	1	MAGIAKLAKDREAAGLGSHERRAIKYINQDYALRNECLEAGTLFPDPSFALPSALGF	60		
DB	1	MAGIAIKLAKDREAAGLGSHERRAIKYINQDYETLRNECLEAGLFPDPSFALPSISLGY	60		
QY	61	KELGYSSKTRGMKRKRPTEICADPOFIIGATRPDIQCALGDGWLAAIASLTLNBEI	120		
DB	61	KELGYSSKTRGIEKMRPTEICADPOFIIGATRPDIQCALGDGWLAAIASLTLNBEI	120		
QY	121	LARVVPILNOSPENAVAGIFHPQPMQYGEVWEVVDRLPTKQGEILFYHSAEGSEFMSAL	180		
DB	121	LARVVPPOQSPENAVAGIFHPQPMQYGEVWEVVDRLPTKQGEILFYHSAEGSEFMSAL	180		
QY	181	LEKAVAKINGCYEALSGGATTEGFEDFTGGIAEWYELKKRPNNLFKIIQKALQKSLGCG	240		
DB	181	LEKAVAKINGCYEALSGGATTEGFEDFTGGIAEWYELKRPNNLFKIIQKALEKSLGCG	240		
QY	241	SIDITSADSEKITEQKLYKHGAYSVTGAEBVESNGSLQKLRIRNPMGEVWGTGRMND	300		
DB	241	SIDITSADSEKITEQKLYKHGAYSVTGAEBVESNGSLQKLRIRNPMQVEMWTGRMND	300		
QY	301	CPSNWTIDPEEERELTRRHEDGFPMWSFSDFLRHSRLIECNLTPDYLTSPTYKKMKLTK	360		
DB	301	CPSNWTVDPEVRANLTERODEGFPMWSFSDFLRHSRLIECNLTPDYLTCDSYKKMKLTK	360		
QY	361	MDGNWRSGSTAGCCNYPNTFMWNPQYILKLEBEDDEDEBGSCTFLVGLIQKARRQR	420		
DB	361	MDGNWRSGSTAGCCNYPNTFMWNPQYILKLEBEDDEDEBGSCTFLVGLIQKARRQR	420		
QY	421	KMGEDMTTIGRCIYVPEBELSGQINIHLSKNEFLTNRAERSDTITNLREVLNRFKLPG	480		
DB	421	KMGEDMTTIGRCIYVPEBELSGQINIHLSKNEFLTNRAERSDTITNLREVLNRFKLPG	480		
QY	481	EYLIIVSPFPEPKDDPFCIRVSEKKADQAVDDLEANTLEEFDSISEDDIDGVRRLPQ	540		
DB	481	EYLIIVSPFPEPKDDPFCIRVSEKKADQAVDDLEANTLEEFDSISEDDIDGVRRLPQ	540		
QY	541	LAGEBAEISAFELQTLIRRLVLAKRODIKSDGFSIEYCKIMVMDLSDSGSLGLKEFYLL	600		
DB	541	LAGEBAEISAFELQTLIRRLVLAKRQDIKSDGFSIEYCKIMVMDLSDSGSLGLKEFYLL	600		
QY	601	WTIKIOKYKIYREIVDNGSTWNSYEMRKALKEAGFKPCQLHQYIVARFADQDILIDPD	660		
DB	601	WTIKIOKYKIYREIVDNGSTWNSYEMRKALKEAGFKPCQLHQYIVARFADDELIDPD	660		
QY	661	NFVRCLVRLTEFLFKIFKQLDPENTGTHIELDLSMCFSTL	700		
DB	661	NFVRCLVRLTEFLFKIFKQLDPENTGTHIELDLSMCFSTL	700		
RESULT 3					
AC	CAN2_RAT	STANDARD;	PRT;	700 AA.	
AC	Q07009;				
DT	01-JUN-1994	(Rel. 29, Created)			
DT	01-JUN-1994	(Rel. 29, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Calpain 2, large [catalytic] subunit precursor (EC 3.4.22.17)				
DE	(Calcium-activated neutral proteinase) (CANP) (M-type) (M-calpain)				
DE	(Molluscan-calpain).				
GN	CANP2.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				

RN SEQUENCE FROM N.A.
 RP MEDLINE=94032492; PubMed=8218419;
 RA Deluca C.I., Davies P.L., Samis J.A., Elce J.S.;
 RT "Molecular cloning and bacterial expression of cDNA for rat calpain
 RT I1 80 kDa subunit";
 RL Biochim. Biophys. Acta 1216:81-93(1993).
 RN [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=21240297; PubMed=11342050;
 RA Moldoveanu T., Hosifield C.M., Jia Z., Elce J.S., Davies P.L.;
 RT "Ca(2+)-induced structural changes in rat m-calpain revealed by
 RT partial proteolysis";
 RL Biochim. Biophys. Acta 1545:245-254(2001).
 RN [3]
 RP MOTAGENESIS OF LYS-230, LYS-234 AND GLUT-504.
 RX MEDLINE=21269273; PubMed=11102442;
 RA Hosifield C.M., Moldoveanu T., Davies P.L., Elce J.S., Jia Z.;
 RT "Calpain mutants with increased Ca2+ sensitivity and implications for
 RT the role of the C(2)-like domain.";
 RL J. Biol. Chem. 276:7404-7407(2001).
 RN [4]
 RP MOTAGENESIS OF CYS-105, HIS-262; ASN-286 AND TRP-288.
 RX MEDLINE=95361909; PubMed=7635186;
 RA Arthur J.S., Gauthier S., Elce J.S.;
 RT "Active site residues in m-calpain: identification by site-directed
 RT mutagenesis.";
 RL FEBS Lett. 368:397-400(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=20069318; PubMed=10601010;
 RA Hosifield C.M., Elce J.S., Davies P.L., Jia Z.;
 RT "Crystal structure of calpain reveals the structural basis for
 RT Ca(2+)-dependent protease activity and a novel mode of enzyme
 RT activation";
 RL EMBO J. 18:6880-6889(1999).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodelling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by 200-1000 micromolar concentrations
 CC of calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding.
 CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, L09120, AAA16327.1, -.
 DR PIR, S38361, S38361.
 DR PDB, IDFO, 26-NOV-01.
 DR MEROPS, C02.002, -.
 DR InterPro, IPR002048, EF-hand.
 DR InterPro, IPR001300, Protease_C2.
 DR InterPro, IPR000169, SHprot_acsite.
 DR Pfam, PF01067, Calpain_I1, 1.
 DR Pfam, PF00036, ehand, 3.
 DR Pfam, PF00648, Peptidase_C2, 1.
 DR PRINTS, PR00704, CALPAIN.
 DR SMART, SM00720, calpain_I1, 1.
 DR SMART, SM00230, CySpC, 1.
 DR SMART, SM00054, Efn, 2.

DR SMART; SMO0054; Efn. 3.
 DR PROSITE; PS00018; EF HAND; 2.
 DR PROSITE; PS00139; THIOL PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL PROTEASE HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL PROTEASE ASN; FALSE NEG.
 KW Hydroxylase; thiol protease; Calcium-binding; Multigene family.
 FT PROPEP 1 19 ANCHORS TO THE SMALL SUBUNIT (POTENTIAL).
 FT CHAIN 20 700 CALPAIN 2, LARGE (CATALYTIC) SUBUNIT.
 FT DOMAIN 20 355 CALPAIN.
 FT DOMAIN 356 514 CALPAIN III.
 FT DOMAIN 515 529 LINKER.
 FT DOMAIN 530 700 DOMAIN IV.
 FT CA BIND 541 552 EF-HAND 1.
 FT CA BIND 585 596 EF-HAND 2.
 FT CA BIND 615 626 EF-HAND 3.
 FT DOMAIN 650 661 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT ACT_SITE 680 691 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 FT ACT_SITE 105 105 BY SIMILARITY.
 FT ACT_SITE 262 262 BY SIMILARITY.
 FT ACT_SITE 286 286 BY SIMILARITY.
 SQ SEQUENCE 700 AA; 79228 MW; C3AEDB39CCB56D3B CRC64;

Query Match 83.4%; Score 3098; DB 1; Length 700;

Best Local Similarity 81.1%; Pred. No. 2.2e-193;

Matches 568; Conservative 67; Mismatches 65; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKREAEGLGSHERAIKYLNODYEALRNCEBAGTLFODSPFPAISALGF 60
 DB 1 MAGIAAKLAKREAEGLGSHERAIKYLNODYEALRNCEBAGTLFODSPFPAISALGF 60
 QY 61 KELGPYSSKTRGMKRPTEICADPOIIGATRTDIOGALGDCWLLAASLTLNEEI 120
 DB 61 RELGPNSYKTKGVWVCRPTLSCPRFAGATRTDIOGALGDCWLLAASLTLNEEI 120
 QY 121 LARVPLNOSQENYAGIFHFQFQYQGWVWVDDRLPTDQGLLPHSAGESEFMSAL 180
 DB 121 LARVPLNOSQENYAGIFHFQFQYQGWVWVDDRLPTDQGLLPHSAGESEFMSAL 180
 QY 121 LARVPLNOSQENYAGIFHFQFQYQGWVWVDDRLPTDQGLLPHSAGESEFMSAL 180
 DB 121 LARVPLNOSQENYAGIFHFQFQYQGWVWVDDRLPTDQGLLPHSAGESEFMSAL 180
 QY 181 LEKAVAKINGCYEALSGATGEGFDFGIAEWYELKKPPNFKIIOKALOGSLLGC 240
 DB 181 LEKAVAKINGCYEALSGATGEGFDFGIAEWYELKKPPNFKIIOKALOGSLLGC 240
 QY 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRWNN 300
 DB 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRWNN 300
 QY 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRWNN 300
 DB 241 SIDTTSADSEAITFOKLKVAHAYSVTGAEEVNSGLQKILIRINPGEVEMTGRWNN 300
 QY 301 CPSNMTIDPEERERKLRHEDGEMFMSDYLKHSLEICNLPPDTITSTYTKKMLTK 360
 DB 301 CPSNMTIDPEERERKLRHEDGEMFMSDYLKHSLEICNLPPDTITSTYTKKMLTK 360
 QY 301 CPSNMTIDPEERERKLRHEDGEMFMSDYLKHSLEICNLPPDTITSTYTKKMLTK 360
 DB 301 CPSNMTIDPEERERKLRHEDGEMFMSDYLKHSLEICNLPPDTITSTYTKKMLTK 360
 QY 361 MDGMMRQSTAGGCRNYPNTFMNPOVYIKLEEDDEDEDESGGCTFLVGLIOKRRROR 420
 DB 361 MDGMMRQSTAGGCRNYPNTFMNPOVYIKLEEDDEDEDESGGCTFLVGLIOKRRROR 420
 QY 361 MDGMMRQSTAGGCRNYPNTFMNPOVYIKLEEDDEDEDESGGCTFLVGLIOKRRROR 420
 DB 361 MDGMMRQSTAGGCRNYPNTFMNPOVYIKLEEDDEDEDESGGCTFLVGLIOKRRROR 420
 QY 421 KMGDMHTIGFIVPEBELSGQTNHLSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 DB 421 KMGDMHTIGFIVPEBELSGQTNHLSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 QY 421 KMGDMHTIGFIVPEBELSGQTNHLSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 DB 421 KMGDMHTIGFIVPEBELSGQTNHLSKNFFLTNRARSDTFINLREVLNFKLPPG 480
 QY 481 EYILVPSFTEPNKGGPCIRVPESEKKADYQAVDEIEANLEPFIDSEDDIDGVARLPAQ 540
 DB 481 EYILVPSFTEPNKGGPCIRVPESEKKADYQAVDEIEANLEPFIDSEDDIDGVARLPAQ 540
 QY 481 EYILVPSFTEPNKGGPCIRVPESEKKADYQAVDEIEANLEPFIDSEDDIDGVARLPAQ 540
 DB 481 EYILVPSFTEPNKGGPCIRVPESEKKADYQAVDEIEANLEPFIDSEDDIDGVARLPAQ 540
 QY 541 LAGEDAEISAFELQTLIRVLAKRODISDGSFISITCINWMDLSDSGKLGKEFYIL 600
 DB 541 LAGEDAEISAFELQTLIRVLAKRODISDGSFISITCINWMDLSDSGKLGKEFYIL 600
 QY 541 LAGEDAEISAFELQTLIRVLAKRODISDGSFISITCINWMDLSDSGKLGKEFYIL 600
 DB 541 LAGEDAEISAFELQTLIRVLAKRODISDGSFISITCINWMDLSDSGKLGKEFYIL 600
 QY 601 WTAKIOKQIYREIDVDRSGTMSYEMKALEAGFKMPCOLHQYVAFADQULITDFD 660
 DB 601 WTAKIOKQIYREIDVDRSGTMSYEMKALEAGFKMPCOLHQYVAFADQULITDFD 660
 QY 601 WTAKIOKQIYREIDVDRSGTMSYEMKALEAGFKMPCOLHQYVAFADQULITDFD 660
 DB 601 WTAKIOKQIYREIDVDRSGTMSYEMKALEAGFKMPCOLHQYVAFADQULITDFD 660
 QY 661 NFRCLVLTLETLFKIKOLDPBNTGTITLILISWICSVL 700
 DB 661 NFRCLVLTLETLFKIKOLDPBNTGTITLILISWICSVL 700
 QY 661 NFRCLVLTLETLFKIKOLDPBNTGTITLILISWICSVL 700
 DB 661 NFRCLVLTLETLFKIKOLDPBNTGTITLILISWICSVL 700

RESULT 5

ID	CANX_CHICK	STANDARD;	PRT;	705 AA.
AC	P00789;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Calpain, large [catalytic] subunit (EC 3.4.22.17) (Calcium-			
DE	activated neutral proteinase) (CAMP) (Mu/M-type).			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85061606; PubMed=6095110;			
RA	Ono S., Emori Y., Imaizumi S., Kawasaki H., Kizaregi M., Suzuki K.;			
RT	"Evolutionary origin of a calcium-dependent protease by fusion of			
RT	genes for a thiol protease and a calcium-binding protein?";			
RL	Nature 312:566-570(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86082358; PubMed=3000828;			
RA	Emori Y., Ono S., Tobita M., Suzuki K.;			
RT	"Gene structure of calcium-dependent protease retains the ancestral			
RT	organization of the calcium-binding protein gene.";			
RL	FEBS Lett. 194:249-252(1986).			
RN	[3]			
RP	CALCIUM-BINDING DATA.			
RX	MEDLINE=87279982; PubMed=3038855;			
RA	Mnam Y., Emori Y., Kawasaki H., Suzuki K.;			
RT	"E-F hand structure-domain of calcium-activated neutral protease			
RT	(CAMP) can bind Ca ²⁺ ions.";			
RL	J. Biochem. 101:889-895(1987).			
RN	[4]			
RP	CHARACTERIZATION.			
RX	MEDLINE=95260862; PubMed=7742367;			
RA	Suzuki K., Tsukihara T., Okada-Ban M., Sugita H., Ishiura S.,			
RT	"Identification of a third ubiquitous calpain species -- chicken			
RT	muscle expresses four distinct calpains.";			
RL	Biochem. Biophys. Acta 1261:381-393(1995).			
CC	-1 FUNCTION: Calcium-regulated non-lysosomal thiol-protease which			
CC	catalyze limited proteolysis of substrates involved in			
CC	cytoskeletal remodeling and signal transduction.			
CC	-1 CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or			
CC	Arg-Xaa with Leu or Val as the P2 residue.			
CC	-1 COFACTOR: Binds 3 calcium ions.			
CC	-1 ENZYME REGULATION: Activated by micromolar concentrations of			
CC	calcium and inhibited by calpastatin.			
CC	-1 SUBUNIT: Heterodimer of large (catalytic) and a small (regulatory)			
CC	subunit.			
CC	-1 SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma			
CC	membrane upon Ca ²⁺ binding (By similarity).			
CC	-1 TISSUE SPECIFICITY: Ubiquitously expressed.			
CC	-1 PTM: THE N-TERMINUS IS BLOCKED.			
CC	-1 SIMILARITY: Contains 5 EF-hand calcium-binding domains.			
CC	-1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.			
CC	-1 CAUTION: THIS PROTEIN WAS PREVIOUSLY THOUGHT TO BE M-CALPAIN BUT			
CC	HAS SINCE BEEN FOUND TO BE AN INTERMEDIATE FORM BETWEEN THE M AND			
CC	MD TYPES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			

EMBL, X01415; CAA25658.1; --
 DR F1R; A00979; CICH.
 DR HSP; Q07009; IDFO.
 DR MEROPS; C02_003; --
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR00169; SH3prot_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 2.
 DR Pfam; PF00648; peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyspc; 1.
 DR SMART; SM00054; Eph; 2.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
 DR Hydrolase; Thiol protease; Calcium-binding; Multigene family.
 KW DOMAIN; 23 358
 FT DOMAIN; 359 517
 FT DOMAIN; 518 533
 FT DOMAIN; 534 704
 FT CA_BIND; 545 556
 FT CA_BIND; 589 600
 FT CA_BIND; 619 630
 FT DOMAIN; 654 665
 FT DOMAIN; 684 695
 FT ACT_SITE; 108 108
 FT ACT_SITE; 265 265
 FT ACT_SITE; 289 289
 SQ SEQUENCE 705 AA; 80351 MW; ABCDCS6298848AA CRC64;
 Query Match 69.5%; Score 2579.5; DB 1; Length 705;
 Best local similarity 66.6%; Pred. No. 8.9e-160;
 Matches 465; Conservative 111; Mismatches 121; Indels 1; Gaps 1;

3 GIAAKLAKDREAAGLGSHERAIKYNQDYALRNECLAEATLFDPSFPAISALGFE 62
 6 GIAARLQDRRLRAVGVEHNNAVYLNQDYALKECTESGTLFRDPQFPAISALGFE 65
 63 LGYSSSTRGMRKRPETICADPOFITGATRTDTCGALGDCMLAIAISLTINEITL 122
 66 LGYSSSTRGMRKRPETICADPOFITGATRTDTCGALGDCMLAIAISLTINEITL 125
 123 RVPPLNOSFOENYAGIFHPQMGVEVEVVDRLPTKCGELLFVHSASGSEFWSLLE 182
 126 RVPHGQSFQEDYAGIFHPQMGVEVVDVDDLLPTKCGELLFVHSASCTEFWSALL 185
 183 KAVAKINGCYEALSGATTEGFEFTGIAEMYLEKKPPNLFKIQIALOKSILGCSI 242
 186 KAVAKINGCYEALSGATTEGFEFTGIAEMYLEKKPPNLFKIQIALOKSILGCSI 245
 243 DITSADSEAITQKLVKHAIVSYTGAEEVNSGLQKLRIRINPWGEVMTGRMNDNC 302
 246 DITSADSEAITQKLVKHAIVSYTGAEEVNSGLQKLRIRINPWGEVMTGRMNDNC 305
 303 SMTTIDEEERBLTRHREDEEFMMSFSDFLRHSRLICNLTPDITLSDTYKKKTKMD 362
 306 EMDNIDSDREELQKMEDEEFMMSFSDFLRHSRLICNLTPDITLSDTYKKKTKMD 365
 363 GNNRSGTAGGCRNYPPTFMNPOYLKLEEDDEDESGGCTFLVGLIOKRRRRORM 422
 366 GNNRSGTAGGCRNYPPTFMNPOYLKLEEDDEDESGGCTFLVGLIOKRRRRORM 425
 423 GEDMTTIGQIYVEBELSGQTNHLSKNPFLTRARERSDTFINREVLNRPFLPGEX 482
 426 GGMHTTIGFAVVEPEBAQGSQVHLTKDFLRQSRARSETFINLEVENQRLRPGEX 485
 483 ILVSTPEPKDGFRCFVSEKKADQAVDDEEAL-DEPDSEDDIDQVRRRLPAOL 541
 486 ILVSTPEPKDGFRCFVSEKKADQAVDDEEAL-DEPDSEDDIDQVRRRLPAOL 545

QY 542 AGEDAIISAELOTITIRRIYAKRQDIKSDGFSIETCKINWMDLSDSGKLGKEFYIWM 601
 DB 546 AGEDAIISAELOTITIRRIYAKRQDIKSDGFSIETCKINWMDLSDSGKLGKEFYIWM 605
 QY 602 TKIQKQKYRELDVDRSGTMSYEMRKALKEGFGPCQLHGVYIARFAPDDLIIDFDN 661
 DB 606 NKIRSWLTTFRQYDLKSGTMSYEMRKALKEGFGPCQLHGVYIARFAPDDLIIDFDN 665
 QY 662 FVACLVKLETFEFTFKPKQDPENTGTIELDISMCFESV 699
 DB 666 FVACLVKLETFEFTFKPKQDPENTGTIELDISMCFESV 703
 RESULT 6
 CAN1_HUMAN STANDARD; PRT; 714 AA.
 ID CAN1_HUMAN
 AC P07384;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated neutral proteinase) (CANP) (Mu-type) (mucanp) (micromolar-calpain).
 GN CAPN1 OR CAPNPL.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86301172; PubMed=3017764;
 RA Aoki K., Imajob S., Ohno S., Emori Y., Koike M., Koseki G., Suzuki K.;
 RT "Complete amino acid sequence of the large subunit of the low-Ca2+-requiring form of human Ca2+-activated neutral protease (mucanp) deduced from its cDNA sequence.";
 RT ERBS Lett. 205:313-317(1986).
 RL [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90380278; PubMed=2400579;
 RA Sorimachi H., Ohmi S., Emori Y., Kawasaki H., Saido T.C., Ohno S., Minami Y., Suzuki K.;
 RT "A novel member of the calcium-dependent cysteine protease family.";
 RT Biol. Chem. Hoppe-Seyler 371:171-176(1990).
 RL [3]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heisch F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ueda T.B., Toshyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bonifard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rauterfeld Y.S.N., Krzywiński M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which catalyze limited proteolysis of substrates involved in cytoskeletal remodeling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or Arg|-Xaa with Leu or Val as the P2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of

CC calcium and inhibited by calpastatin.
 CC -1 SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
 CC membrane upon Ca++ binding (By similarity).
 CC -1 SIMILARITY: Contains 5 EF-hand calcium-binding domains.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04366; CAA27881.1; -
 DR EMBL; BC008751; AAH08751.1; -
 DR PIR; A26213; C1H0H.
 DR HSSP; Q07009; IDPO.
 DR MEROPS; C02.001; -
 DR Genew: HGNC:1476; CAPNL.
 DR MIM; 114220; -
 DR GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease_C2.
 DR InterPro; IPR00169; SHprot acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00036; ehand; 3.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; Cyspc; 1.
 DR SMART; SM00054; Efh; 3.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 KM Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
 FT DOMAIN 30 365 CALPAIN
 FT DOMAIN 365 526 DOMAIN III
 FT DOMAIN 527 542 LINKER
 FT DOMAIN 543 713 DOMAIN IV
 FT CA_BIND 554 565 EF_HAND 1.
 FT CA_BIND 598 609 EF_HAND 2.
 FT CA_BIND 628 639 EF_HAND 3.
 FT DOMAIN 663 674 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT DOMAIN 693 704 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
 FT ACT_SITE 115 115 By similarity.
 FT ACT_SITE 272 272 By similarity.
 FT ACT_SITE 296 296 By similarity.
 FT CONFLICT 548 548 K -> N (IN REF. 3).
 SQ SEQUENCE 714 AA; 81889 MW; 1CB6D7C56D063498 CRC64;

Query Match 66.2%; Score 2458.5; DB 1; Length 714;
 Best local similarity 63.1%; Pred. No. 6.3e-152;
 Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

QY 3 GIAKLAKDRRAAGLSGSHERAIKYLNODYALNRECEAGTLQDPSFPALPALKGEK 62
 DB 13 GVSAAVOQARRELGRLGHENAIRKLGQDYQLRRCQSGTLPRDEAFPPVPSOLGKX 72
 QY 63 LGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDQCMWLAIAISLTINEELIA 122
 DB 73 LGRNPSKTYGIRKMRPTILSNPOFIYDGAFTDICGALDQCMWLAIAISLTINDTLIH 132
 QY 123 RVLPLNOSPENYAGIFHFQFQWQYGEWVVDRLPTKDELFLVHSAGSEFWSALLE 182
 DB 133 RVVPHGSGFQNGVAGIFHFQFQWQYGEWVVDRLPTKDELFLVHSAGSEFWSALLE 192
 QY 183 KAYKINGCYALSGGATTEFEDFTGIAWYELKPPPLFKIKQALOKSLGCSI 242
 DB 193 KAYKINGCYALSGGATTEFEDFTGIAWYELKPPPLFKIKQALOKSLGCSI 252

QY 243 DITSADSEAITFOKLKVGNAVSVTGAEVENSGLQKILIRINPWGEVMTGRNDNCP 302
 DB 253 DISSVLMEMAITFKLVKGNAVSVTGAKQNVYGVLSLRMNPGEVMTGAMDSSS 312
 QY 303 SMTIDPEERERLTRRHEDGEFWMFSDFIRHYSRLICNLIPDITLSDTYKKWKLTKMD 362
 DB 313 EANNVDPYERDQLRVKMEDEPFWMSPDFRFRFTRLICNLITDALKSRIRKMTLLYE 372
 QY 363 GNNRSGTAGCCNNYPTFMNPOYLKLEENDEDE--DGSGCTFLVLQKHRROR 420
 DB 373 GYNRSGTAGCCNNYPTFMNPOYLKLEENDEDE--DGSGCTFLVLQKHRROR 432
 QY 421 KMGEDMTITFGIYVEPEELSGQTNHLSKNFELTNRRARSDPTINLRVLRFXLP 480
 DB 433 RFRDMETIGFAYVEVPELVGPAYHLKRDFFLANASRRSQFINLRVSTRFLPLPG 492
 QY 481 EYLLVSTPEPNTGDFCIRVSEKADYQAVDEIBANL--EFPDISEDIDGVRRLPA 539
 DB 493 EYVVPSTPEPNTGDFCIRVSEKADYQAVDEIBANL--EFPDISEDIDGVRRLPA 552
 QY 540 QLAGEDEIAFELQTLIRVYLAKRQDISGSEISCTKIMVMDLSDSGKGLKEFYI 599
 DB 553 QLAGEDEIAFELQTLIRVYLAKRQDISGSEISCTKIMVMDLSDSGKGLKEFYI 612
 QY 600 LMTKIQKYKRYEIDVDKSGTNNVSEMRKALEEAGFKMPCQHQIYVARFADQILDF 659
 DB 613 LMRIRNYLSIFKPLDLSGSMAYEMRAIISAGFKLKKYELLITRYSPLDANVF 672
 QY 660 DNFVRCIVLETLFKIPKQIDPENTGTIELDISWL 695
 DB 673 DNFVRCIVLETLFKIPKQIDPENTGTIELDISWL 708

RESULT 7
 ID CANL_PIG STANDARD; PRT; 714 AA.
 AC P35750; Q29600; O9N0M6;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 1, large (catalytic) subunit (EC 3.4.22.17) (Calcium-activated
 DE neutral proteinase) (CANP) (Mu-type) (mucanp) (micromolar-calpain).
 GN CAPNL.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP Smith T.P.L., Simmen F.A., Vallet J.A.;
 RT "Rapid Communication: nucleotide sequences of two isoforms of porcine
 RT micromolar calcium-activated neutral proteinase 1 (mucalpain) cDNA";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 326-415 FROM N.A.
 RT TISSUE=Small intestine;
 RA Winteroe A.K., Fredholm M., Davies W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 528-623 FROM N.A.
 RT TISSUE=skeletal muscle;
 RC MEDLINE=9416155; Pubmed=8312396;
 RA Sun W., Ji S.Q., Ebert P.J., Bitwell C.A., Hancock D.L.;
 RT "Cloning the partial cDNAs of mu-calpain and m-calpain from porcine
 RT skeletal muscle.";
 RL Biochimie 75:931-936(1993).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
 CC catalyze limited proteolysis of substrates involved in
 CC cytoskeletal remodelling and signal transduction.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the p2 residue.
 CC -1- COFACTOR: Binds 3 calcium ions (By similarity).

```

CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC membrane upon Ca++ binding (by similarity).
CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF263610; AAF3444.1; -.
CC EMBL; F14611; CAA23154.1; -.
CC EMBL; U01180; AAA65125.1; -.
CC HSSP; Q07009; IDFO.
CC MEROPS; C02.001; -.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR001300; Protease_C2.
CC InterPro; IPR001069; Shprot_acsite.
CC Pfam; PF01067; Calpain_III; 1.
CC Pfam; PF00036; ehand; 3.
CC Pfam; PF00648; Peptidase_C2; 1.
CC PRINTS; PR00704; CALPAIN.
CC SMART; SM00720; calpain_III; 1.
CC SMART; SM00230; Cyspc; 1.
CC SMART; SM00054; Efn; 3.
CC PROSITE; PS00018; EF_HAND; 2.
CC PROSITE; PS00139; THIOL_PROTEASE_HIS; 1.
CC PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
CC PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
CC HydroLase; Thiol; protease; Calcium-binding; Multigene family.
CC -----
CC FT DOMAIN 30 365
CC FT DOMAIN 366 526
CC FT DOMAIN 527 542
CC FT DOMAIN 543 713
CC FT CA_BIND 554 565
CC FT CA_BIND 598 609
CC FT CA_BIND 628 639
CC FT DOMAIN 663 704
CC FT DOMAIN 693 704
CC FT ACT_SITE 115 115
CC FT ACT_SITE 272 272
CC FT ACT_SITE 296 296
CC FT CONFLICT 528 528
CC FT CONFLICT 531 531
CC FT CONFLICT 541 541
CC FT CONFLICT 622 622
CC SEQUENCE 714 AA; 81738 MW; 0BB31DBEFC56369A CRC64;

Query Match 65.8%; Score 2444.5; DB 1; Length 714;
Best Local Similarity 63.1%; Pred. No. 5, 1e-151;
Matches 440; Conservative 120; Mismatches 132; Indels 5; Gaps 3;

```

```

Db 193 KAVAKNGSYEALSGSGTSEGFEDFTGVTWELRKADSDVLSILKALENGSLIGCSI 252
Qy 243 DITSADSEATITFOKLVKGNAYSVTGAEEVNSGSLQKLRIRNPMGEVMTGRANDNC 302
Db 253 DISVADMEAVTTKLVKGNAYSVTGAQVNVGGVNLIRNPMGEVMTGAMDGSS 312
Qy 303 SMTTIDPEERERTLRHEDGEFMSFSDPLRHYSRLICNLTPDTLTSPTYKKMKLTAMD 362
Db 313 ENMGVPYORDQLRVMEDEGEFMSFSDPLRHYSRLICNLTPDTLTSPTYKKMKLTAMD 372
Qy 363 GMRBRSSTAGCGCNVNTFMNPOYLKLEEDDEDE--GSGCTPLVGLIQKRRRO 419
Db 373 GWRKRSSTAGCGCNVNTFMNPOYLKLEEDDEDE--GSGCTPLVGLIQKRRRR 432
Qy 420 RKMGEMHTITGFIYVEPELSCQTNIHLSKNPLFNRAERSDTFINTREVLNFKLP 479
Db 433 RFRGRMETITGFIYVEPELSCQTNIHLSKNPLFNRAERSDTFINTREVLNFKLP 491
Qy 480 GEYILVSTPEPNKDDFCIRVSEKKADYQAVDEIEANT--BEFDSIEDDIDGVRRLP 538
Db 492 GEYVVPSTPEPNKDDFCIRVSEKKADYQAVDEIEANT--BEFDSIEDDIDGVRRLP 551
Qy 539 AQLAGDARIISAFELQTLIRVLAQRQDIKSDFSLETCKINVMILDSGSGTGLTKERY 598
Db 552 RQLAGDMEISVARELTITNRIISKHKDLRTKGFSLSCRSVMNIMDRDGNGLGLVERV 611
Qy 599 ILWTIKQYOKTYREIDVDRSGTMSYEMRKALSEAGFMPQOLHOVIARFADPDLLID 658
Db 612 ILMNRIRNVLISFRKDLDSGMSAYEMRMALIESGFLNKKLFELITRISPEPLAVD 671
Qy 659 FDNFVRCVLRLETLFRKDLDPENTGTELDLSWL 695
Db 672 FDNFVRCVLRLETLFRKDLDPENTGTELDLSWL 708

RESULT 8
CAN1 MOUSE STANDARD; PRT; 713 AA.
ID CAN1_MOUSE 088666;
AC 035350; 088666;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calpain 1, large (catalytic) subunit (EC 3.4.22.17) (Calcium-activated
GN neutral proteinase) (CANP) (Mu-type) (mucANP) (micromolar-calpain).
OS CAPN1 OR CANP1 OR CAPN1.
OC Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CF-1;
RA Poirier C., Poussard S., Faust D.M., Imaizumi-Sherer T., Weiss M.C.,
RA Ducastaing A., Montarras D., Pinset C., Guenet J.-L.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RA Sahr K.E., Andrabi S., Peters L.L., Chishti A.H.;
RT "Cloning and characterization of the cDNA and gene encoding the mouse
RT mu-calpain large subunit protein.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyze limited proteolysis of substrates involved in
CC cytoskeletal remodeling and signal transduction.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
CC Arg-Xaa with leu or val as the p2 residue.
CC -1- COFACTOR: Binds 3 calcium ions (by similarity).
CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma

```

```

CC membrane upon Ca++ binding (By similarity).
CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF021847; AAB72222.1; -
DR EMBL: AF084459; AAC33134.1; -
DR HSSP: Q07009; IDPO.
DR MEROPS: C02.001; -.
DR MGD; MGI:88263; Capn1.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR001300; Protease C2.
DR InterPro: IPR000169; SHprot acstce.
DR Pfam: PF01067; Calpain_III; 1.
DR Pfam: PF00036; ehand; 3.
DR Pfam: PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; Cyspc; 1.
DR SMART; SM00054; Eph; 3.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
KW Hydrolase; Thiol protease; Calcium-binding; Multigene family.
FT DOMAIN 30 365 CALPAIN
FT DOMAIN 366 525 DOMAIN III.
FT DOMAIN 526 541 LINER.
FT DOMAIN 542 712 DOMAIN IV.
FT CA BIND 553 564 EF-HAND 1.
FT CA BIND 597 608 EF-HAND 2.
FT CA BIND 627 638 EF-HAND 3.
FT DOMAIN 662 673 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT DOMAIN 692 703 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
FT ACT_SITE 115 115 BY SIMILARITY.
FT ACT_SITE 272 272 BY SIMILARITY.
FT ACT_SITE 296 296 BY SIMILARITY.
FT CONFLICT 488 488 L -> P (IN REF. 2).
FT CONFLICT 696 696 D -> N (IN REF. 2).
SQ SEQUENCE 713 AA; 82106 MW; 3E1E26C95802B864 CRC64;
Query Match 65.6%; Score 2435; DB 1; Length 713;
Best Local Similarity 62.8%; Pred. No. 2,1e-150;
Matches 437; Conservative 126; Mismatches 129; Indels 4; Gaps 3;

```

```

DB 313 ENKVDVPEREQARVAKMEDDEPFMMSPRDFRREPTKLEICNLTPDLAKSRTLRMNTTFYE 372
QY 363 GNNRSGTAGGCNNYNTFMNPOYLKLEED- EDEEDSGGCTFLVGLQKRROR 420
DB 373 GTRRSTAGGCNNYATFVNPQKIRLEEVADADYDNRGGCSFLALQKRRER 432
QY 421 KMGEDHTTGFGIYVEPEELSGQTNTHLSKNFPLTNRARERSPTFNLREVLRFPLPG 480
DB 433 RFRGDDMETIGFANYQVPRELAGP-VHLKRDFFLANASQSHFNLNRVSRIRLPG 491
QY 481 EYILVSTPEEPKNDGFCIRVFSEKKADYQAVDEIEANL-EEPDISEDIDIDGVRLLPA 539
DB 492 EYILVSTPEEPKNDGFCIRVFSEKKADYQAVDEIEANL-EEPDISEDIDIDGVRLLPA 551
QY 540 QLAGDDEAIEAFLOTLIRVLAARODIKSDGSIETCKMVMMDSDSGKGLKEFYI 599
DB 552 KLGGDDMEISVKELOTLIRVLAARODIKSDGSIETCKMVMMDSDSGKGLKEFYI 611
QY 600 LMTKIOYKIVREIVDRSGTNSYEMRKALEEAGFKMPQCHQYIARFADQILIDF 659
DB 612 LMKRIRNYLTIFPKFDLDSGMSAYEMRAIAGFKMLKHLITRYSPLDANVF 671
QY 660 DNEFRCIVRELTFLKIFKOLDPENTGTIELDLISWL 695
DB 672 DNEFRCIVRELTFLKIFKOLDPENTGTIELDLISWL 707
RESULT 9
CANT RAT STANDARD; PRT; 713 AA.
ID CANT RAT
AC P9571;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Calpain 1, large [catalytic] subunit (EC 3.4.22.17) (Calcium-activated
DE neutral proteinase) (CANP) (Mu-type) (mucANP) (Micromolar-calpain).
GN CAPN1 OR CLS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97107433; Pubmed=8950173;
RA Sorimachi H., Amano S., Ishiura S., Suzuki K.;
RT "Primary sequences of rat mu-calpain large and small subunits are,
RT respectively, moderately and highly similar to those of human.";
RL Biochim. Biophys. Acta 1309:37-41(1996).
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyze limited proteolysis of substrates involved in
CC cytoskeletal remodeling and signal transduction.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
CC Arg-Xaa with Leu or Val as the P2 residue.
CC -1- COFACTOR: Binds 3 calcium ions (By similarity).
CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
CC calcium and inhibited by calpastatin.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic; Translocates to the plasma
CC membrane upon Ca++ binding (By similarity).
CC -1- SIMILARITY: Contains 5 EF-hand calcium-binding domains.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U53858; AAC53001.1; -

```


Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00054; Efn; 3.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS_PARTIAL.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS_PARTIAL.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN_PARTIAL.
KM Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
FT NON_TER 1 1
FT DOMAIN 1 77 CALPAIN
FT DOMAIN 78 236 DOMAIN III.
FT DOMAIN 237 251 LINKER.
FT DOMAIN 252 422 DOMAIN IV.
FT CA_BIND 263 274 EF_HAND 1.
FT CA_BIND 307 318 EF_HAND 2.
FT CA_BIND 337 348 EF_HAND 3.
FT DOMAIN 372 383 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT DOMAIN 402 413 ANCESTRAL CALCIUM SITE 5 (POTENTIAL).
FT ACT_SITE 8 8 BY SIMILARITY.
SQ SEQUENCE 422 AA; 49494 MW; AE4FA3C48A33C41 CRC64;
Query Match 58.0%; Score 2153; DB 1; Length 422;
Best Local Similarity 93.8%; Pred. No. 2e-132;
Matches 396; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
QY 279 OKLRINPMGEVMTGRMNDNCSPWNTIDPEEREERLTRHEDGPFMMSPDFLRHYSRL 338
DB 1 OKLRINPMGEVMTGRMNDNCSPWNTIDPEEREERLAEHEDGPFMMSPDFLRHYSRL 60
QY 339 EICMLPTDTLSDTYKKMKLTKMDGNMRGSGTAGCGRNYPNTFMWNPOLYIKLEEDDE 398
DB 61 EICMLPTDTLSDTYKKMKLTKMDGNMRGSGTAGCGRNYPNTFMWNPOLYIKLEEDDEQ 120
QY 399 EDGSGCTFLVGLQKRRRQRKKGEDMHTGFGIYERPEELSCQTINHLKSNFPLNRA 458
DB 121 EDGSGCTFLVGLQKRRRQRKKGEDMHTGFGIYERPEELSCQTINHLKSNFPLNRA 180
QY 459 RERSDTFLNREVNRFPLPGEYILVSTPEPKDGFRCFVSEKKADYQAVDEIEA 518
DB 181 RERSDTFLNREVNRFPLPGEYILVSTPEPKDGFRCFVSEKKADYQAVDEIEA 240
QY 519 NLEEFDISDDIDDGVRRLFQAQAGEPAISAFELQTLIRKVLAKRODIKSGFSIETCK 578
DB 241 NLEEFDISDDIDDGVRRLFQAQAGEPAISAFELQTLIRKVLAKRODIKSGFSIETCK 300
QY 579 IMVMDLSDSGSGKGLKEFYILMTKIQYQKIYREIDVDRSGTMSYEMRKALBEAGFKM 638
DB 301 IMVMDLSDSGSGKGLKEFYILMTKIQYQKIYREIDVDRSGTMSYEMRKALBEAGFKL 360
QY 639 PCQGHQIVARFADDOILIDFNFRCLVRELETFKIPIKOLDPENTGTIELDLISWLCFS 698
DB 361 PCQGHQIVARFADDOILIDFNFRCLVRELETFKIPIKOLDPENTGTIELDLISWLCFS 420
QY 699 VL 700
DB 421 VL 422

RESULT 11

CANB HUMAN STANDARD; PRT; 702 AA.
AC 09U06;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calpain 11 (EC 3.4.22.17) (Calcium-activated neutral proteinase 11)
DE (CANP 11).
GN CAPN11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_Taxid=9606;

RN [1]
RX SEQUENCE FROM N.A. PubMed=10409436;
RP MEDLINE=9933989;
RA Dear T.N., Moller A., Boehm T.;
RT "CAPN11: A calpain with high mRNA levels in testis and located on
RT chromosome 6."
RL Genomics 59:243-247(1999).
CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease which
CC catalyze limited proteolysis of substrates involved in
CC cytoskeletal remodelling and signal transduction.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or
CC Arg|-Xaa with Leu or Val as the P2 residue.
CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
CC (regulatory) subunit.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN TESTIS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ242832; CAB52137.1; -.
DR HSSP; Q07009; 1DFO.
DR MEROPS; C02.013; -.
DR GENEW; HGNC:1478; CAPN11.
DR MIM; 604822; -.
DR GO; GO:0004198; F:calpain activity; TAS.
DR INTERPRO; IPR002048; EF-hand.
DR INTERPRO; IPR001300; Protease_C2.
DR INTERPRO; IPR001659; SHPTC_acsite.
DR Pfam; PF01067; Calpain_III; 1.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00648; Peptidase_C2; 1.
DR PRINTS; PR00704; CALPAIN.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; Cyapc; 1.
DR SMART; SM00054; Efn; 2.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 2.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
KM Hydrolyase; Thiol protease; Calcium-binding; Multigene family.
FT DOMAIN 1 77 CALPAIN
FT DOMAIN 78 236 DOMAIN III.
FT DOMAIN 237 251 LINKER.
FT DOMAIN 252 422 DOMAIN IV.
FT CA_BIND 263 274 EF_HAND 1.
FT CA_BIND 307 318 EF_HAND 2.
FT CA_BIND 337 348 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
FT ACT_SITE 102 102 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
FT ACT_SITE 259 259 BY SIMILARITY.
FT ACT_SITE 283 283 BY SIMILARITY.
SQ SEQUENCE 702 AA; 80582 MW; 3A61862B9674FAE CRC64;
Query Match 54.2%; Score 2014; DB 1; Length 702;
Best Local Similarity 52.4%; Pred. No. 4e-123;
Matches 365; Conservative 133; Mismatches 193; Indels 4; Gaps 2;
QY 4 IAAKLADRFAGLGSHERAITYLADYALNCECLAEFTLFQDSFPALPAGLKFEL 63
DB 1 MVAHINNSRLKAGVGQHDVAQNFQNSFEELFAACLRKELDEDDPLFPAEPLSLGKDL 60
QY 64 GPYSKTRGMRKRPKEICADPOFIIGGARTRTICGAGDCKLAAISLTINLETIAR 123
DB 61 GPNSKVVQNTSWQRPDIINPLFIMDGISPTDICGILADCKLAAISLTICPKILYR 120
QY 124 VVPLNQSFGQENVAGIPIHFQFWQGEWVEVVDRLPTKQGLFLVHSAGSGRFWSALLEK 183

Db 121 VPRGSGFKKNGVIGIFHFOIMQFGQWVNVVDRLPTKNDKLVFVHSTESSEFMSALIEK 180
 QY 184 AYAKINGCYEALSGATTEGFEDEFTGGIAEMVELEKPPPLPTIOALOKSGILGCSID 243
 Db 181 AYAKSGSYALSGSGMELEDEFTGGVAVSFOLOPPHLLKILRAVRSLSMGCSIE 240
 QY 244 ITSAADSEALTPOKLVKGAHAYVTGAEEVNSLSQKILIRNPWEVETGRMNDNCP 303
 Db 241 VTSDESELSWTDKMLVKGHAYVTGLQDVHGRGMETLIRVNPWGRIENNGAMSPARE 300
 QY 304 WNTIDBERERLRRRHEDGFPMSPDFLHYSRLICNLTPTLTISDYTKMKMLTKMDG 363
 Db 301 MEESVADICQWOLHAKHEDGFPMSPDYODFLNNFTLLEICNLTPTLTISDYTKMKMLTKMDG 360
 QY 364 NMRGSGTAGGCRVYPTFMNPOVLLKLEBEDDEDESGS---CPFLVGLIOKHRROR 420
 Db 361 SMRSGSAGGCRNHPGFTWNPQKISLPEGDDPEDABGNVAVCTLVLMQKRNHAR 420
 QY 421 KMGEDMHTIGFGIYEVEPEELSGQNTIHLKSNFPLTNRARSSTFNLREVLNRFKLP 480
 Db 421 OQGAOQOTIGFVLYAVPKEFQNTQDVHLKKEPFTKYGDHGFSEIFTNRSRVSSQLRLPG 480
 QY 481 EYILVSTPEPNKDGPCIFVSEKADYAVD-ELEANLEPFDSIEDIDGVRRLPA 539
 Db 481 EYILVSTPEPNKDGPCIFVSEKADYAVD-ELEANLEPFDSIEDIDGVRRLPA 539
 QY 540 QLAGEDAEISAFELQITIRVLAKRODICKSDFSIETCKIMVMLDSGSGKLGKFEYI 599
 Db 541 IVAGEKEIGVVELQRLNMAIKFSPKTKGFGLDRCMIMNDKSGKLGKLEFKI 600
 QY 600 LMTKIOKYOKIYEIDVRSQTMNSYEMKALEAGFKMPCOLHOVIVARFADOLIIDF 659
 Db 601 LMKLKKMMDIFRECQDHSGLTNSYEMRLVIEKAGIKLNNKMQVLYAVAYADDLIIDF 660
 QY 660 DNVRCLVLETLFKIQLDPENTGTIELDISWL 695
 Db 661 DSFISCFRLKLTWTFEFLTWDPKNTGHCILSLEOWL 696
 RESULT 12
 ID CAN3_CHICK STANDARD; PRT; 810 AA.
 AC 092177;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 3 large subunit (EC 3.4.22.17) (Calpain I3) (Calpain p94,
 DE large [catalytic] subunit) (Calcium-activated neutral proteinase 3)
 DE (CANP 3) (Muscle-specific calcium-activated neutral protease 3 large
 DE subunit).
 GN CAPN3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=lung;
 RX MEDLINE=95260862; PubMed=7742367;
 RA Sorimachi H., Tsukahara T., Okada-Ban M., Sugita H., Ishiura S.,
 RA Suzuki K.;
 RT "Identification of a third ubiquitous calpain species -- chicken
 RT muscle expresses four distinct calpains";
 RL Biochim. Biophys. Acta 1261:381-393(1995).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with leu or val as the P2 residue.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (potential).
 CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE. LOW LEVELS IN SPLEEN,
 CC INTESTINE AND BONE.
 CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D38028; BAA07230.1; -;
 DR PIR; S57196; S57196.
 DR HSRP; Q07009; IDP0.
 DR MEROPS; C02.004; -;
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR001300; Protease C2.
 DR InterPro; IPR00169; SHPoc. acsite.
 DR Pfam; PF01067; Calpain III; 1.
 DR Pfam; PF00036; ehand_4; 4.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR PRINTS; PR00704; CALPAIN.
 DR SMART; SM00720; calpain III; 1.
 DR SMART; SM00230; CysPc; 1.
 DR SMART; SM00054; Efb; 3.
 DR PROSITE; PS0018; EF_HAND; 2.
 DR PROSITE; PS00139; THIOL PROTEASE CYS; 1.
 DR PROSITE; PS00639; THIOL PROTEASE HIS; FALSE NEG.
 DR PROSITE; PS00640; THIOL PROTEASE ASN; FALSE NEG.
 KW Hydrolyase; Thiol protease; Calcium-binding; Nuclear protein;
 KW Multi-gene family.
 FT DOMAIN 43 421
 FT DOMAIN 432 579
 FT DOMAIN 580 638
 FT DOMAIN 639 809
 FT CA_BIND 694 705
 FT CA_BIND 724 735
 FT DOMAIN 759 770
 FT DOMAIN 789 800
 FT ACT_SITE 123 123
 FT ACT_SITE 327 327
 FT ACT_SITE 351 351
 SQ
 Query Match 810 AA; 93560 MW; E8DE99411C51041E CRC64;
 Best Local Similarity 51.1%; Score 1897; DB 1; Length 810;
 Matches 350; Conservative 145; Mismatches 177; Indels 98; Gaps 7;
 QY 18 LGSHERAIKYNODYALNRNECLEAGTLFQDPSEFPAIPBALGKKEIGPSSKTRGRWR 77
 Db 41 ISNQPIIKKKEKTYEELHKKLEENILVEDDPFPNENSLFYSQKVPK-----PEWR 95
 QY 78 PTEICADPOFIIGATRTDICGALGDCWLLAIASTLTNEETLAVVPLNOSFOENYAG 137
 Db 96 PREICENPRFIIGANRTDICGELDQCFALAIACITLANKLLCIVTHIDQSFIQYAG 155
 QY 138 IFHFQFQYGEWVNVVDRLPTKDELLFVHSAGSEFMSALIEKAYAKINGCYEALSG 197
 Db 156 IFHFQFQYGEWVNVVDRLPTKDELLFVHSAGSEFMSALIEKAYAKINGCYEALSG 215
 QY 198 GATTEGFEDEFTGGIAEMVELEKPPPLPTIOALOKSGILGCSID-----243
 Db 216 GNTTEMEFTGGVIVFYEIKDAPKDIYKIMKRIARGSLMSIDNIGFHYGAAPRS 275
 QY 244 -----ITSAADSEALTPOKLVKGAHAYVTGAEEVNSLSQKILIRNPWEVETGRMNDNCP 270
 Db 276 IGLIARMKVLENAQMTSYVDYQSTDERPANTIMPQYETRMSCGLVKGHAYSTYAV 335
 QY 271 EYVNSLSQKILIRNPWEVETGRMNDNCPSEWNTIDBERERLRR-HEDEGFMSFS 329

QY 258 -----LYKGAYSYTGAEVESNSLOKLR 283
 DB 296 NSLLRSDSLDPRGSDDRPSRTIVVOYETRMACGLVKGHAYSVTGLEALFKGEKVLVR 355
 QY 284 IRRPMGEVETGWMNDCPSMNTIDPEERELTRR-HEDEPFMSFSDFLRHSRLKICN 342
 DB 356 LRRPMQGVENWNGSMDKMSFVDKDEKARLOHQVTEDEFWMSYDDFYHFTKLEICN 415
 QY 343 LTPDRLTSDTYKKWKLTKMDGNRRGSTAGCCNRYPTFMNPOYLKLEEDDEDEGE 402
 DB 416 LTPDALLESDELQMTWVSVEGRRVKGCSAGCCNRPFTFTNTNPOYKLEEDDDPDSSE 475
 QY 403 SGGTFLVGLIOKRRRRQKNGEDMTTIGFQIYEVPEELSGQTNHLKNEFLTNRARERS 462
 DB 476 VICSFLVALMOKRRKRRKLGANLFTIGFAIYEPKEMHGNKO-HLQKDFLVNASKRS 534
 QY 463 DTINLRVLRKLPGEITLVSTPEPNKDDFCIRVSEKK-----506
 DB 535 KTYNNKEVSGRRPLPSEYIVPSTYEPHOGEFFILRVSEKRNLSBAENTISYDRPV 594
 QY 507 -----ADYQAVDEIEANLEB-----FDSIEDDIDGV 534
 DB 595 KKKKKPITIVSDRANSKELGVDOAEKGRKAGPEKGETPOPRGHTDQSEBEOQOF 654
 QY 535 RLPLAQLAGEDAISAFELQTIILRRVLAKRQDIKSDGFSIETCKIVMDLSDSGGLGL 594
 DB 655 RNIFRQAGDMEICADELKNVINTVYVNGKDKLTQGFTESSCRSMALMDTQSSGLNL 714
 QY 595 KEYIILMTKIQKQYKIREIDVDRSGTMSYENRKALEEFGKPCOLHOVIYARPDQ 654
 DB 715 QEFHMLMKKIKAMOKIFKHYDTDHSCTINSYEMNANVNDAGFHLNSQDYIITWRVADKH 774
 QY 655 LIIDPFRVACLVRLTETLFKIFKQIDPEENTGTELDLSML 695
 DB 775 MNIDPDSFICFVRLBSMFRAFNAPKDGITIKLANLEWL 815

RESULT 14
 ID CAN3_RAT STANDARD; PRT; 821 AA.
 AC P16259;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Calpain 3 large subunit (BC 3.4.22.17) (Calpain I3) (Calpain p94,
 DE large [catalytic] subunit) (Calcium-activated neutral proteinase 3)
 DE (CANP 3) (Muscle-specific calcium-activated neutral protease 3 large
 DE subunit).
 GN CAPN3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Skeletal muscle;
 RX MEDLINE=90062125; PubMed=2555341;
 RA Sorimachi H., Imaizoh-Omori S., Emori Y., Kawasaki H., Ohno S.,
 RA Minami Y., Suzuki K.;
 RT "Molecular cloning of a novel mammalian calcium-dependent protease
 RT distinct from both m- and mu-types. Specific expression of the mRNA
 RT in skeletal muscle.";
 RL J. Biol. Chem. 264:20106-20111(1989).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-Xaa, Met-Xaa or
 CC Arg-Xaa with Leu or Val as the P2 residue.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Skeletal muscle.

CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 CC EMBL: J05121; AAA41790.1; --
 DR PIR: B34488; B34488.
 DR HESP: Q07009; IDPO.
 DR MEROPS: C02.004; --
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR001300; Protease_C2.
 DR InterPro: IPR000169; SHPoc_acsite.
 DR Pfam: PF01067; Calpain_I1; 1.
 DR Pfam: PF00036; ehfand; 3.
 DR Pfam: PF00648; Peptidase_C2; 1.
 DR PRINTS: PR00704; CALPAIN_1.
 DR SMART: SM00720; calpain_I1; 1.
 DR SMART: SM00230; CyPpc; 1.
 DR SMART: SM00054; Efn; 3.
 DR PROSITE: PS00018; EF_HAND; 2.
 DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.
 DR PROSITE: PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 DR PROSITE: PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 DR Hydrolase; Thiol protease; Calcium-binding; Multigene family; Repeat.
 FT DOMAIN 49 428
 FT DOMAIN 429 586
 FT DOMAIN 587 649
 FT DOMAIN 650 820
 FT CA_BIND 705 716
 FT CA_BIND 735 746
 FT DOMAIN 770 781
 FT DOMAIN 800 811
 FT ACT_SITE 129 129
 FT ACT_SITE 334 334
 FT ACT_SITE 358 358
 SQ SEQUENCE 821 AA; 94127 MW; 27FAEAD2FEA19FBF CRC64;
 Query Match 50.4%; Score 1869.5; DB 1; Length 821;
 Best local Similarity 46.1%; Pred. No. 1.2e-113;
 Matches 351; Conservative 131; Mismatches 176; Indels 103; Gaps 8;
 QY 32 YEALNRECELAGTLFODPSFPALPSALGFKELEGYSSKTRGMKRPTEICADPOFTIGG 91
 DB 61 FEQLHKKCLEKVKLVLDPEPDPETSLFYSQKEPIQ-----FWKKPPEICENPRLTIGG 115
 QY 92 ATRTDTCGALGECWMLAIASTLNEELIARVPLNOSFOEYVAGIFHFQOYGEWVE 151
 DB 116 ANRTDTCGDLGECWMLAIACTLNERLLFRVTPHDSQTEVEYAGIFHFQYRVDWD 175
 QY 152 VVVDRLPTKDGELLFVHSAEGSEFWSALLEKAYAKINGCYBALSGATTEGDEDTGFI 211
 DB 176 VVDDDLPTNNOLVTKSHRNEFPSSALLEKAYALHSGSYELKKGNTTEAMEDTGGV 225
 QY 212 AEWEELKPEPNLFKTIQKALQKSLGCSID-----ITSAADSEAITFOK-----257
 DB 236 TEFPEIKDAVSDWYKIMRAIERGSLMGCSIDGTMTVTGTPSGLMGELIARVRNMD 295
 QY 258 -----LYKGAYSYTGAEVESNSLOKLR 283
 DB 296 NSLLRSDSLDPRASDRPSRTIVVOYETRMACGLVKGHAYSVTGLEALFKGEKVLVR 355
 QY 284 IRRPMGEVETGWMNDCPSMNTIDPEERELTRR-HEDEPFMSFSDFLRHSRLKICN 342
 DB 356 LRRPMQGVENWNGSMDKMSFVDKDEKARLOHQVTEDEFWMSYDDFYHFTKLEICN 415
 QY 343 LTPDRLTSDTYKKWKLTKMDGNRRGSTAGCCNRYPTFMNPOYLKLEEDDEDEGE 402

Db 416 LTADALBESDKLTQTVTSVNEGRVWRCGAGCGRNFPDTFWTNPOYRLKLTLEEDDDPPDSE 475
 Qy 403 SCGTEFLVGLIQKRROROKMGEDMHTTGFYIYEPBELSGOTNHLKSNFLITWRARERS 462
 Db 476 VTCSPFLVAMOKNRKDKRKLGNLFTTICFALYEPKKEHGNKQ-HLQKDFLYNASKARS 534
 Qy 463 DTFINLRVNLRFKLPGEYILVSTPEPNKDGFCIRVFSEKK----- 506
 Db 535 KTYINMEVSGRFLPPESEYIVPSTYEPHGEFILKVFSEKNLSEANENTISVDRPV 594
 Qy 507 -----ADYQ-----ANDEIFRANLE-----FDISEDDIDDGV 534
 Db 595 KKKKKKPIIFVSDRANSKELGVQDEAEKGKTKGPKQSGSPORPGHTDQSESEQQCF 654
 Qy 535 RRLFAOLAGEPAEISAFELQTLIRRVLAKRODITSGDSGFIECTKIMVMDLSDSGKGL 594
 Db 655 RNTFRQIAGDMELCABDLKAVLNTVYVKKHDKLTQGTSLSCSMILMNTDSSGRNL 714
 Qy 595 KEFYILWTKIQYOKIYREIDVDSGTMNSYEMRKALBEAGFKMPCQLHOVIVARFADQ 654
 Db 715 QEFHLWKKIKAMQKIFGHYDTHSGTINSYEMRAVADGPHLNSQLYDITWRVADKH 774
 Qy 655 LIIPDNVRLVRLTETFKIFKQIDPENTGTIELDISWL 695
 Db 775 MNIDFDSFICCFVRLGEMFRAFHAFDKDGIIRKLNVLEWL 815

RESULT 15

CAN3 HUMAN STANDARD; PRT; 821 AA.
 AC P20807; Q9Y5S6; Q9Y5S7;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calpain 3 large subunit (EC 3.4.22.17) (Calpain I2) (Calpain p94,
 large [calpain] subunit) (Calcium-activated neutral proteinase 3)
 DE (CAMP 3) (Muscle-specific calcium-activated neutral protease 3 large
 subunit) (ncr-1).
 GN CAPN3 OR CAPN3 OR CAPN3 OR NCCL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND VARIANTS LGMD2A.
 RX MEDLINE=95236448; PubMed=7720071;
 RA Richard I., Broux O., Allmand V., Fougereuse F.,
 Chatain M., Roudaut C., Hillaire D., Passos-Bueno M.-R., Zatz M.,
 Tischerfeld J.A., Fardieu M., Jackson C.E., Cohen D., Beckmann J.S.;
 RT "Mutations in the proteolytic enzyme calpain 3 cause limb-girdle
 muscular dystrophy type 2A.";
 RL Cell 81:27-40(1995).
 RN [2]
 RP SEQUENCE OF 44-821 FROM N.A.
 RX MEDLINE=90062125; PubMed=2555341;
 RA Sorimachi H., Imajoh-Ohmi S., Emori Y., Kawasaki H., Ohno S.,
 Mitani Y., Suzuki K.;
 RT "Molecular cloning of a novel mammalian calcium-dependent protease
 distinct from both m- and mu-types. Specific expression of the mRNA
 in skeletal muscle.";
 RL J. Biol. Chem. 264:20106-20111(1989).
 RN [3]
 RP SEQUENCE FROM N.A. (ISORFMS II AND III).
 RA Dickson J.M.U., Love D., Evans C.W.E.;
 RT "Alternatively exon-spliced isoforms of calpain 3 expressed in human
 leukocytes.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP REVIEW ON VARIANTS:
 RX MEDLINE=99264234; PubMed=10330340;
 RA Richard I., Roudaut C., Saenz A., Pogue R., Grimbergen J.E.M.A.,
 Anderson L.V.B., Beley C., Cobo A.-M., de Diego C., Eymard B.,

RA Gallano P., Ginjaar H.B., Laea A., Pollitt C., Topaloglu H.,
 Rutzberger J.A., de Visser M., van der Kooi A., Bushby K., Bakker E.,
 Lopez de Munain A., Fardieu M., Beckmann J.S.;
 RT "Calpainopathy -- a survey of mutations and polymorphisms.";
 RL Am. J. Hum. Genet. 64:1524-1540(1999).
 RN [5]
 RP VARIANTS LGMD2A GLN-572 AND GLY-744.
 RX MEDLINE=96176567; PubMed=8624690;
 RA Fardieu M., Hillaire D., Mignard D., Mignard J., Feingold N., Feingold J.,
 Mignard D., de Ubeda B., Collin H., Tome F.M.S., Richard I.,
 Beckmann J.S.;
 RT "Juvenile limb-girdle muscular dystrophy. Clinical, histopathological
 and genetic data from a small community living in the Reunion
 Island.";
 RL Brain 119:295-308(1996).
 RN [6]
 RP VARIANTS LGMD2A.
 RX MEDLINE=97294404; PubMed=9150160;
 RA Richard I., Broux O., Dincer P., Roudaut C., Bady B.,
 Burgunder J.-M., Chemaly R., Garcia C.A., Halaby G., Jackson C.E.,
 Kureit D.M., Lefranc G., Legum C., Loiselet J., Merlin L.,
 Niveleon-Chevallier A., Ollagnon-Roman E., Restagno G., Topaloglu H.,
 Beckmann J.S.;
 RT "Multiple independent molecular etiology for limb-girdle muscular
 dystrophy type 2A patients from various geographical origins.";
 RL Am. J. Hum. Genet. 60:1128-1138(1997).
 RN [7]
 RP VARIANTS LGMD2A ASN-336; GLN-490; VAL-702 AND GLN-748.
 RX MEDLINE=97410018; PubMed=9266733;
 RA Dincer P., Leturcq F., Richard I., Piccolo F., Yalinoglu D.,
 de Toma C., Akcoeren Z., Broux O., Deburtgrave N., Bengtler L.,
 Roudaut C., Rutzberger J.A., Jung D., Tan E., Jeanpierre M.,
 Campbell K.P., Kaplan J.-C., Beckmann J.S., Topaloglu H.;
 RT "A biochemical, genetic, and clinical survey of autosomal recessive
 limb girdle muscular dystrophies in Turkey.";
 RL Hum. Neurol. 42:222-229(1997).
 RN [8]
 RP VARIANTS LGMD2A ARG-222; GLU-486; TRP-489 AND GLN-748.
 RX MEDLINE=98433925; PubMed=9762961;
 RA Urtasun M., Saenz A., Roudaut C., Poza J.J., Rutzberger J.A.,
 Cobo A.-M., Richard I., Garcia Bragado F., Leturcq F., Kaplan J.-C.,
 Marti Maeso J.F., Beckmann J.S., Lopez de Munain A.;
 RT "Limb-girdle muscular dystrophy in Guipuzcoa (Baeque Country,
 Spain).";
 RL Brain 121:1735-1747(1998).
 RN [9]
 RP VARIANT LGMD2A 200-PHE-LEU-204 DEL.
 RX MEDLINE=98112493; PubMed=9452114;
 RA Haeffner K., Speer A., Huebner C., Volt T., Cexle K.;
 RT "A small in-frame deletion within the protease domain of muscle-
 specific calpain, p94 causes early-onset limb-girdle muscular
 dystrophy 2A.";
 RL Hum. Mutat. Suppl. 1:5298-S300(1998).
 RN [10]
 RP VARIANT LGMD2A GLY-744.
 RX MEDLINE=98317835; PubMed=9655129;
 RA Penisson-Besnier I., Richard I., Dubas F., Beckmann J.S., Fardieu M.;
 RT "Pseudometabolic expression and phenotypic variability of calpain
 deficiency in two siblings.";
 RL Muscle Nerve 21:1078-1080(1998).
 RN [11]
 RP VARIANT LGMD2A CYS-360.
 RX MEDLINE=98442618; PubMed=9771675;
 RA Kawai H., Akaike M., Kunishige M., Inui T., Adachi K., Kimura C.,
 Kawai Y., Nishida Y., Endo I., Kashwagi S., Nishino H.,
 Fujiwara T., Okuno S., Roudaut C., Richard I., Beckmann J.S.,
 Miyoshi K., Matsumoto T.;
 RT "Clinical, pathological, and genetic features of limb-girdle muscular
 dystrophy type 2A with new calpain 3 gene mutations in seven patients
 from three Japanese families.";
 RL Muscle Nerve 21:1493-1501(1998).
 CC -1- FUNCTION: Calcium-regulated non-lysosomal thiol-protease.
 -1- CATALYTIC ACTIVITY: Preferential cleavage: Tyr|-Xaa, Met|-Xaa or

CC Arg-|-Xaa with Leu or Val as the P2 residue.
 CC -1- ENZYME REGULATION: Activated by micromolar concentrations of
 CC calcium and inhibited by calpastatin.
 CC -1- SUBUNIT: Heterodimer of a large (catalytic) and a small
 CC (regulatory) subunit.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=3;
 CC Name=1;
 CC IsoId=P20807-1; Sequence=Displayed;
 CC Name=1;
 CC IsoId=P20807-2; Sequence=VSP_005227, VSP_005228;
 CC Name=1;
 CC IsoId=P20807-3; Sequence=VSP_005229;
 CC -1- TISSUE SPECIFICITY: Skeletal muscle.
 CC -1- DISEASE: Defects in CAPN3 are the cause of limb-girdle muscular
 CC dystrophy type 2A (LGM2A) [MIM:253600]. LGM2A is both autosomal
 CC dominantly and recessively transmitted. It is characterized by
 CC progressive symmetrical atrophy and weakness of the proximal limb
 CC muscles and elevated serum creatine kinase. The symptoms usually
 CC begin during the first two decades of life, and the disease
 CC gradually worsens, often resulting in loss of walking ability 10
 CC or 20 years after onset.
 CC -1- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C2.
 CC -1- DATABASE: NAME=Leiden Muscular Dystrophy pages;
 CC NOTE=Calpain-3 mutations in LGM2A;
 CC WWW=http://www.dmd.nl/capn3_home.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X85030; CAA59403.1; -
 CC EMBL; AF127764; AAD28253.1; -
 CC EMBL; AF127765; AAD28254.3; -
 CC PIR; A56218; CIH03.
 CC HSRP; Q07009; IDP0.
 CC MEROPS; C02.004; -
 CC DR GeneW; HGNC:1480; CAPN3.
 CC DR MIM; 114240; -
 CC DR MIM; 253600; -
 CC DR GO; GO:0005622; C:intracellular; TAS.
 CC DR GO; GO:0004198; F:calpain activity; TAS.
 CC DR GO; GO:0008234; F:cysteine-type peptidase activity; TAS.
 CC DR GO; GO:0004871; F:signal transducer activity; TAS.
 CC DR GO; GO:0007517; F:muscle development; TAS.
 CC DR InterPro; IPR002048; EF-hand.
 CC DR InterPro; IPR001300; Protease_C2.
 CC DR InterPro; IPR000169; SHPoc_acsite.
 CC DR Pfam; PF01067; Calpain_III; 1.
 CC DR Pfam; PF00648; Peptidase_C2; 1.
 CC DR PRINTS; PR00704; CALPAIN_III; 1.
 CC DR SMART; SM00720; calpain_III; 1.
 CC DR SMART; SM00230; Cyspc; 1.
 CC DR SMART; SM00054; Efn; 3.
 CC DR PROSITE; PS00018; EF_HAND; 2.
 CC DR PROSITE; PS00139; THIOL_PROTEASE_CYS; 1.
 CC DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE NEG.
 CC DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE NEG.
 CC DR HydroLase; Thiol protease; Calcium-binding; Multigene family; Repeat;
 CC Disease mutation; Polymorphism; Alternative splicing.
 CC FT DOMAIN 49 428 CALPAIN.
 CC FT DOMAIN 429 586 DOMAIN III.
 CC FT DOMAIN 587 649 LINKER.
 CC FT DOMAIN 650 820 DOMAIN IV.
 CC FT CA_BIND 705 716 EF-HAND 1 (PROBABLE).
 CC FT CA_BIND 735 746 EF-HAND 2 (PROBABLE).

FT DOMAIN 770 781 ANCESTRAL CALCIUM SITE 3 (POTENTIAL).
 FT DOMAIN 800 811 ANCESTRAL CALCIUM SITE 4 (POTENTIAL).
 FT ACT_SITE 129 129 BY SIMILARITY.
 FT ACT_SITE 334 334 BY SIMILARITY.
 FT ACT_SITE 358 358 BY SIMILARITY.
 FT VARSPIC 268 315 Missing (in isoform II).
 FT VARSPIC 595 639 /FTId=VSP_005227.
 FT VARSPIC 595 639 /FTId=VSP_005228.
 FT VARSPIC 595 639 /FTId=VSP_005228.
 Query Match 50.2%; Score 1862.5; DB 1; length 821;
 Best Local Similarity 45.5%; Pred. No. 3.3e-113;
 Matches 347; Conservative 139; Mismatches 171; Indels 105; Gaps 8;
 QY 32 YEALRNECEAGLTPDPSFPALPSALGFELOPYSKTRGMKRPTEICADPOFIIG 91
 DB 61 FEQLHKKCLEKLYLVDPPEPPEPDESLFSQKPEI0---FWKKEPPECNPRFTIDG 115
 QY 92 ATRTDICOGALGDCWLLAAIASLTINBEILARVVPINQSPQENYAGIFHPQMGWE 151
 DB 116 ANRTDICOSELGDCWFLAALACLTINQHLFRVPHDQSTIENYAGIFHPQMGWE 175
 QY 152 VVVDRLPTKQGLLFVHSAEGSEFMSALLKAVYKINGCYEALSGGATTEGFEPTGTI 211
 DB 176 VVIDDCLPTYNQGLVFTKSNHRNRPMSALLEKAYAKHGSEYALKGNTTEAMEFTGV 235
 QY 212 AEWYELKKPPPNFLFKIIOKRALOGSLIGSID----ITSABEALTFQK----- 257
 DB 236 AEFELIRADSDWYKIMKKAIERGSLMGCSIDGTNTGTSPSGNMGELIARMVRND 295
 QY 258 -----LVKGAHYSTGAEEVESNSLOKILR 283
 DB 296 NSLLQSDLDPRGSDPRPTITIPQYETRMAGLVRGHAIVSTGLDVEYFKKXKLV 355
 QY 284 IRNPGEVETGWKNDNCBSWNTIDPEBERLTR- HEDGEFMSFSDFLRHSRLICN 342
 DB 356 LRNPQGVENWGSWSRDKWSEFVDKDEKARLOHQTVEDEFFMSYEDFLYHFTKLEICN 415
 QY 343 LTPEDTLTSPYKKKTKLTKNDGNRRSTAGGCNNTYPMNQYLKLEEDDEDEGE 402
 DB 416 LTVADALQSDPLQTVTSVVEGRVRCSSAGGCNPFDTFTWNTQYRLKLEEDDDPDS 475
 QY 403 SCCTFLVGLIQKRRRQRKMGEDMTTIGCIYVEPELGGQTIHLSKNPFLTRARERS 462
 DB 476 VICSFLVALMQKRRKDRKLGASLFTTIGRAIVPEKMGKNC-HQKDFLYNASKARS 534
 QY 463 DTFINLREVLNREKLPPEGEYIIVSTFEPKQDFCIRVFSEKKADYQAVDEIEANTL-- 520
 DB 535 KTYINRREVSQRRLPPESEVIAIVSTYEPHQBEPFLRVFSEKRNISEVENTISYDRPV 594
 QY 521 -----BEFDSEDDIDG----- 533
 DB 595 KKKTKPRIIFVSDRANSNKELGVDSE-BEGKGTSPDKQKOPOPQSSDSEBEOQ 653
 QY 534 VRLLPQALGEDAIEAFELQTLRLRVLLAKKODIKSDGSIETCKLWMDLDSGSGKG 593
 DB 654 FRNIFQIAGDDMEICADELKKYLVVYKHKDHLHGFTLSCRSMTALMDTDSGSKLN 713
 QY 594 LKEFYILMTKIOKQKITYEIRIDVRSGTMSYEMKALBEAGKMPCOLHQVIVARFAD 653
 DB 714 LGEFHLNKKIKAMQKIKPHYDIDQGTINSYEMRANVADAGHLNQQYDITIMYADK 773
 QY 654 QLIIDPDNFRCLVRETLFKIKQDPENTGTIEDLISWL 695
 DB 774 HMDIDDSFICCFRLEGMFRAHARDKQDGIKLVLEWL 815

Search completed: July 24, 2003, 12:52:29
 Job time : 21.3504 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:46:04 ; Search time 34.3469 Seconds

(without alignments)
1959.945 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713

Sequence: 1 MAGIAAKLAKDREAEGIGS.....PENTGTIEDLISWLGFSVL 700

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3705	99.8	700	1 C1H0H2	calpain (EC 3.4.22
2	3518	94.7	700	2 S38361	calpain (EC 3.4.22
3	3098	83.4	700	2 S57194	calpain (EC 3.4.22
4	2579.5	69.5	705	1 C1CHH	calpain (EC 3.4.22
5	2458.5	66.2	714	1 C1H0H	calpain (EC 3.4.22
6	2356.5	63.5	703	2 A48764	calpain (EC 3.4.22
7	2153	58.0	422	2 B24815	calpain (EC 3.4.22
8	1984	53.4	586	2 S57195	calpain (EC 3.4.22
9	1897	51.1	810	1 S57196	calpain (EC 3.4.22
10	1869.5	50.4	821	1 B34488	calpain (EC 3.4.22
11	1862.5	50.2	821	1 C1H0H3	calpain (EC 3.4.22
12	1594.5	42.9	805	1 A55054	calpain (EC 3.4.22
13	1199.5	32.3	758	1 A39343	calpain (EC 3.4.22
14	1184	31.9	653	1 S44749	COG64.2 protein -
15	925.5	24.9	320	2 A24815	calpain (EC 3.4.22
16	859.5	22.6	647	2 UC5772	tissue-specific ca
17	789	21.2	648	2 S71885	sex-determining pr
18	508	13.6	2143	2 G96595	hypothetical prote
19	458	12.3	263	2 A34466	calpain (EC 3.4.22
20	458	12.3	266	1 C1PGU	calpain (EC 3.4.22
21	454	12.2	266	1 C1RBL	calpain (EC 3.4.22
22	450	12.1	268	1 C1H0L	calpain (EC 3.4.22
23	435.5	11.7	184	2 A55143	calpain (EC 3.4.22
24	426	11.5	1597	2 BVFFSL	sol protein, large
25	426	11.5	1597	2 T08428	gene small optic 1
26	394	10.6	634	2 T32871	hypothetical prote
27	380	10.2	616	2 T24825	hypothetical prote
28	291	7.8	805	2 T24431	hypothetical prote
29	278	7.5	617	2 T22175	hypothetical prote

30	274	7.4	623	2 T22177	hypothetical prote
31	240	6.5	198	2 S52094	sorcin - human
32	239	6.4	217	2 A42578	EF-hand calcium-bi
33	236	6.4	198	2 A25706	sorcin - Chinese h
34	208.5	5.6	1134	2 C88040	protein P47F6.5 [l
35	188	5.1	829	2 AH2282	hypothetical prote
36	177.5	4.8	425	2 T23122	hypothetical prote
37	171	4.6	611	2 A11941	hypothetical prote
38	168.5	4.5	186	2 D84673	probable calcium b
39	164.5	4.4	153	2 S04970	calcium-binding pr
40	146.5	3.9	335	2 S64352	hypothetical prote
41	146.5	3.9	501	2 AH1852	hypothetical prote
42	143	3.9	482	2 S27608	cysteine proteinas
43	134.5	3.6	217	2 T26132	hypothetical prote
44	127.5	3.4	216	2 G70163	hypothetical prote
45	126.5	3.4	169	1 BCKM	caltracrin - Chlam

ALIGNMENTS

RESULT 1

C1H0H2
calpain (EC 3.4.22.17) large chain 2 - human
N/Alternate names: calpain chain L-2; calpain II catalytic chain; high-calcium requirin
C/Species: Homo sapiens (man)
C/Date: 21-Nov-1993 #sequence revision 09-Aug-1997 #text_change 16-Jul-1999
C/Accession: S10590; A31218; A33529
R/Sorimachi, H.; Ohno, S.; Emori, Y.; Kawasaki, H.; Saido, T.C.; Ohno, S.; Minami, Y.;
BioI. Chem. Hoppe-Seyler 371(Suppl.), 171-176, 1990
A/Title: A novel member of the calcium-dependent cysteine protease family.
A/Reference number: S10589; PMID:90380278; PMID:2400579
A/Accession: S10590
A/Molecule type: mRNA

A/Residues: 1-700 <SOR>
R/Imajoh, S.; Aoki, K.; Ohno, S.; Emori, Y.; Kawasaki, H.; Sugihara, H.; Suzuki, K.
Biochemistry 27, 8122-8128, 1988
A/Title: Molecular cloning of the cDNA for the large subunit of the high-Ca(2+)-requirin
A/Reference number: A31218; PMID:8916474; PMID:2852952
A/Accession: A31218
A/Molecule type: protein
A/Residues: 1-210, 'T', 212-394, 'D', 396-445, 'T', 447-700 <IMA>
A/Cross-references: GB:M23254; NID:G511636; PIDN:AAA5645.1; PID:G511637
A/Note: parts of this sequence were determined by protein sequencing; the amino end of
R/Hata, A.; Ohno, S.; Akita, Y.; Suzuki, K.
U. Biol. Chem. 264, 6404-6411, 1989
A/Title: Tandemly reiterated negative enhancer-like elements regulate transcription of
A/Reference number: A33529; PMID:89197947; PMID:2539381
A/Accession: A33529
A/Molecule type: DNA
A/Residues: 1-67, 'G', 69-72, 'IE', 75-78, 'R', <HAT>
A/Cross-references: DDBJ:J04700; NID:G179910; PIDN:AAA52760.1; PID:G463086.
C/Genetic:
A/Gene: GDB:CAPN2; mCAPN; CAPN1
A/Cross-references: GDB:119750; OMIM:114230
A/Map position: 1pter-1qter
C/Complex: heterodimer of L (large) and S (small) chains
C/Function:
A/Description: catalyzes the hydrolysis of peptides
A/Note: cleaves preferentially after tyrosine, methionine, or arginine residues and bef.
C/Specificity: calpain large chain; calmodulin repeat homology; calpain catalytic domai.
C/Keywords: acetylated amino end; calcium binding; cysteine proteinase; EF hand; hetero.
F:2-700/Product: calpain large chain 2 #status predicted <MAT>
F:75-337/Domain: calpain catalytic domain homology <CLP>
F:529-560/Domain: calmodulin repeat homology <EF1>
F:572-604/Domain: calmodulin repeat homology <EF2>
F:605-634/Domain: calmodulin repeat homology <EF3>
F:637-669/Domain: calmodulin repeat homology <EF4>
F:2/Modified site: acetylated amino end (Ala) (in mature form) #status predicted
F:105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 99.8%; Score 3705; DB 1; Length 700;
Best Local Similarity 99.6%; Pred. No. 3,1e-238;

Matches 697; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MAGIAAKIAXDREAEGSGHERAIAKYINDYBALNCEAGTLFODPSPPAIPSAIGF 60
Db 1 MAGIAAKIAXDREAEGSGHERAIAKYINDYBALNCEAGTLFODPSPPAIPSAIGF 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 120
Db 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 120
Qy 121 LARVPLNOSFOENYAGIFHFQFQYGEWVAVDDRLPTXGDELLFVHSAEGSEFMSAL 180
Db 121 LARVPLNOSFOENYAGIFHFQFQYGEWVAVDDRLPTXGDELLFVHSAEGSEFMSAL 180
Qy 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEMWELKPPNLFKIIQKALOKSILGCG 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEMWELKPPNLFKIIQKALOKSILGCG 240
Qy 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKLIIRINPMGEVEMTGRMND 300
Db 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKLIIRINPMGEVEMTGRMND 300
Qy 301 CPSMNTIDPERERELTRHEDGEPFMSFDFLRHYSRLICNLTPTDITSDTYKKMKLTG 360
Db 301 CPSMNTIDPERERELTRHEDGEPFMSFDFLRHYSRLICNLTPTDITSDTYKKMKLTG 360
Qy 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLILKEEDEDDEDESGCTFLVGLIQKRRROR 420
Db 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLILKEEDEDDEDESGCTFLVGLIQKRRROR 420
Qy 421 KMGEDMTIGFGIYEVPEELSGQTNILSKNFLTNRARSDTFINLREVLNRPKLPBG 480
Db 421 KMGEDMTIGFGIYEVPEELSGQTNILSKNFLTNRARSDTFINLREVLNRPKLPBG 480
Qy 481 EYILVPSFEPHKGDFCIRVSEKKADYQAVDDIEANLEEDISEDDIDGVRRLFAQ 540
Db 481 EYILVPSFEPHKGDFCIRVSEKKADYQAVDDIEANLEEDISEDDIDGVRRLFAQ 540
Qy 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGLKEFYIL 600
Db 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGLKEFYIL 600
Qy 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVIVARFADQIIFD 660
Db 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVIVARFADQIIFD 660
Qy 661 NFVRCVLRLTFLFKIFKQIDPENGTIETLIDLSWLCFSVL 700
Db 661 NFVRCVLRLTFLFKIFKQIDPENGTIETLIDLSWLCFSVL 700

```

RESULT 2

S38361

calpain (EC 3.4.22.17) II large chain - rat

N/Alternate names: calpain II 80k chain

C/Species: Rattus norvegicus (Norway rat)

C/Date: 31-Dec-1993 #sequence revision 02-Aug-1994 #text_change 22-Jun-1999

C/Accession: S38361; S08650; S39751

R/DeLucca, C.I.; Davies, P.L.; Samis, J.A.; Elce, J.S.

Biochim. Biophys. Acta 1216, 81-93, 1993

A/Title: Molecular cloning and bacterial expression of cDNA for rat calpain II 80 kDa subunit

A/Reference number: S38361; MUID:94032492; PMID:8218419

A/Accession: S38361

A/Molecule type: mRNA

A/Residues: 1-700

A/Cross-references: EMBL:109120; NID:9402665; PIDN:AAA16327.1; PID:9402666

R/Samis, J.A.; Back, D.W.; Graham, E.J.; Elce, J.S.

submitted to the EMBL Data Library, February 1990

A/Reference number: S08650

A/Accession: S08650

A/Molecule type: DNA

A/Residues: 380-439 <SAM>

A/Cross-references: EMBL:X51772

C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C/Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyc

F/75-327/Domain: calpain catalytic domain homology <CALP>

F/529-560/Domain: calmodulin repeat homology <EF1>

F/572-604/Domain: calmodulin repeat homology <EF2>

F/605-634/Domain: calmodulin repeat homology <EF3>

F/637-669/Domain: calmodulin repeat homology <EF4>

F/105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 94.7%; Score 3518; DB 2; Length 700;

Beat Local Similarity 93.4%; Pred. No. 8.2e-226; Indels 0; Gaps 0;

Matches 654; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

```

Qy 1 MAGIAAKIAXDREAEGSGHERAIAKYINDYBALNCEAGTLFODPSPPAIPSAIGF 60
Db 1 MAGIAAKIAXDREAEGSGHERAIAKYINDYBALNCEAGTLFODPSPPAIPSAIGF 60
Qy 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 120
Db 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAIATLTNEEI 120
Qy 121 LARVPLNOSFOENYAGIFHFQFQYGEWVAVDDRLPTXGDELLFVHSAEGSEFMSAL 180
Db 121 LARVPLNOSFOENYAGIFHFQFQYGEWVAVDDRLPTXGDELLFVHSAEGSEFMSAL 180
Qy 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEMWELKPPNLFKIIQKALOKSILGCG 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEMWELKPPNLFKIIQKALOKSILGCG 240
Qy 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKLIIRINPMGEVEMTGRMND 300
Db 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKLIIRINPMGEVEMTGRMND 300
Qy 301 CPSMNTIDPERERELTRHEDGEPFMSFDFLRHYSRLICNLTPTDITSDTYKKMKLTG 360
Db 301 CPSMNTIDPERERELTRHEDGEPFMSFDFLRHYSRLICNLTPTDITSDTYKKMKLTG 360
Qy 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLILKEEDEDDEDESGCTFLVGLIQKRRROR 420
Db 361 MDGNMRRGSTAGGCNRYNPTFMNPOYLILKEEDEDDEDESGCTFLVGLIQKRRROR 420
Qy 421 KMGEDMTIGFGIYEVPEELSGQTNILSKNFLTNRARSDTFINLREVLNRPKLPBG 480
Db 421 KMGEDMTIGFGIYEVPEELSGQTNILSKNFLTNRARSDTFINLREVLNRPKLPBG 480
Qy 481 EYILVPSFEPHKGDFCIRVSEKKADYQAVDDIEANLEEDISEDDIDGVRRLFAQ 540
Db 481 EYILVPSFEPHKGDFCIRVSEKKADYQAVDDIEANLEEDISEDDIDGVRRLFAQ 540
Qy 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGLKEFYIL 600
Db 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGLKEFYIL 600
Qy 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVIVARFADQIIFD 660
Db 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVIVARFADQIIFD 660
Qy 661 NFVRCVLRLTFLFKIFKQIDPENGTIETLIDLSWLCFSVL 700
Db 661 NFVRCVLRLTFLFKIFKQIDPENGTIETLIDLSWLCFSVL 700

```

RESULT 3

S57194

calpain (EC 3.4.22.17) large chain 2 - chicken

N/Alternate names: m-calpain heavy chain

C/Species: Gallus gallus (chicken)

C/Date: 28-Oct-1995 #sequence revision 03-Nov-1995 #text_change 20-Jun-2000

C/Accession: S57194

R/Sorimachi, H.; Tsukahara, T.; Okada-Bar, M.; Sugita, H.; Ishiura, S.; Suzuki, K.

Biochim. Biophys. Acta 1261, 381-393, 1995

A/Title: Identification of a third ubiquitous calpain species - chicken muscle expressed

A/Reference number: S57194; MUID:95260862; PMID:7742367

A/Accession: S57194
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-700 <SOR>
 A/Cross-references: EMBL:D8026; NID:9882068; P1DN:BA07228.1; P1D:9882069
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: cysteine protease; EF hand; hydrolase
 F/5-337/Domain: calpain catalytic domain homology <CALP>
 F/529-560/Domain: calmodulin repeat homology <EF1>
 F/572-604/Domain: calmodulin repeat homology <EF2>
 F/605-634/Domain: calmodulin repeat homology <EF3>
 F/637-669/Domain: calmodulin repeat homology <EF4>
 F/670-700/Domain: calmodulin repeat homology <EF5>
 F/105,262,286/Active site: Cys, His, Asn #status predicted

Query Match 83.4%; Score 3098; DB 2; Length 700;
 Best Local Similarity 81.1%; Pred. No. 6,5e-198;
 Matches 568; Conservative 67; Mismatches 65; Indels 0; Gaps 0;

QY 1 MAGIAAKLADREABAGLSHERAIKYNODYEALRNECLEAGTLFODPSFPAIPALGF 60
 DB 1 MAGIAAKLADREABAGLSHERAIKYNODYEALRNECLEAGTLFODPSFPAIPALGF 60
 QY 61 KEIGPYSKTRGMRKRPTEICADPOFTIGATRTDIOGALGDCWLLAALASLTINNEI 120
 DB 61 KEIGPYSKTRGMRKRPTEICADPOFTIGATRTDIOGALGDCWLLAALASLTINNEI 120
 QY 121 LARVPLNOSFOENYAGIFHFOFMQGEWVEVVDLPTKDGELLFVHSAEGSEFMSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFOFMQGEWVEVVDLPTKDGELLFVHSAEGSEFMSAL 180
 QY 181 LERAKVAKINGCYEALSGGATTEGFEFTGGIAEWELKRPENLFKIIQALOKGSLG 240
 DB 181 LERAKVAKINGCYEALSGGATTEGFEFTGGIAEWELKRPENLFKIIQALOKGSLG 240
 QY 241 SIIITSAADSEAITPQKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 300
 DB 241 SIIITSAADSEAITPQKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 300
 QY 241 SIIITSAADSEAITPQKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 300
 DB 241 SIIITSAADSEAITPQKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 300
 QY 301 CPSEMTIDPERERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDILTSDTYKKMKLTK 360
 DB 301 CPSEMTIDPERERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDILTSDTYKKMKLTK 360
 QY 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRRQR 420
 DB 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRRQR 420
 QY 421 KMGEDMHTIGFAIYEVPEPFGQTNIHLKSNFELTNARAKSNFTFIMLREVLNRFKLPAG 480
 DB 421 KMGEDMHTIGFAIYEVPEPFGQTNIHLKSNFELTNARAKSNFTFIMLREVLNRFKLPAG 480
 QY 481 EYIVLSTPEPNDDGDCIRVPESEKADYQAVNDEIRANLEEPDISDDIDDOGRRLFAQ 540
 DB 481 EYIVLSTPEPNDDGDCIRVPESEKADYQAVNDEIRANLEEPDISDDIDDOGRRLFAQ 540
 QY 541 LAEDDAEISAFELQTIIRRLVLAQRDIDSGDSFISCTKIMWMDLSDSGSKLGKEFYIL 600
 DB 541 LAEDDAEISAFELQTIIRRLVLAQRDIDSGDSFISCTKIMWMDLSDSGSKLGKEFYIL 600
 QY 601 WTKIQOKYKIREIDVDRSGTMSYEMRRALAEAGFLYSLQHLQIIVARFADDELIDFD 660
 DB 601 WTKIQOKYKIREIDVDRSGTMSYEMRRALAEAGFLYSLQHLQIIVARFADDELIDFD 660
 QY 661 NPVRCIVRLETLFKIFQDLDPENTGTIELDLSLCSYL 700
 DB 661 NPVRCIVRLETLFKIFQDLDPENTGTIELDLSLCSYL 700
 QY 661 NPVRCIVRLETLFKIFQDLDPENTGTIELDLSLCSYL 700
 DB 661 NPVRCIVRLETLFKIFQDLDPENTGTIELDLSLCSYL 700

RESULT 4
 CICH
 calpain (BC 3.4.22.17) large chain 4 - chicken
 N/Alternate names: calpain catalytic chain; intermediate calcium activated neutral protease
 C/Species: Gallus gallus (chicken)

C/Date: 17-May-1985 #sequence_revision 09-Aug-1997 #extc change 24-Nov-1999
 C/Accession: A00979
 R/Ohno, S.; Emori, Y.; Imajoh, S.; Kawasaki, H.; Kisaragi, M.; Suzuki, K.
 Nature 312, 566-570, 1984
 A/Title: Evolutionary origin of a calcium-dependent protease by fusion of genes for a calmodulin repeat homology and a cysteine protease
 A/Reference number: A93348; MUID:85061606; PMID:6095110
 A/Accession: A00979
 A/Molecule type: mRNA
 A/Residues: 1-705 <OHN>
 A/Cross-references: EMBL:X01415; NID:963332; P1DN:CAA25658.1; P1D:963333
 R/Emori, Y.; Ohno, S.; Tobita, M.; Suzuki, K.
 FEBS Lett. 194, 249-252, 1986
 A/Reference number: A91354; MUID:86082358; PMID:3008628
 A/Contents: annotation; gene structure
 C/Comment: This calpain has calcium requirements intermediate between those of the high and low molecular weight calpains
 C/Genetics:
 A/Insertions: 82/3; 106/1; 145/3; 190/2; 246/3; 274/3; 303/2; 328/2; 382/1; 438/3; 442/3;
 A/Complex: heterodimer of L (large) and S (small) chains
 C/Function:
 A/Description: catalyzes the hydrolysis of peptides
 C/Note: cleaves preferentially after tyrosine, methionine, or arginine residues and before
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: blocked amino end; calcium binding; cysteine protease; duplication; EF hand
 F/5-337/Domain: calpain catalytic domain homology <CALP>
 F/533-564/Domain: calmodulin repeat homology <EF1>
 F/576-608/Domain: calmodulin repeat homology <EF2>
 F/609-638/Domain: calmodulin repeat homology <EF3>
 F/641-673/Domain: calmodulin repeat homology <EF4>
 F/674-705/Domain: calmodulin repeat homology <EF5>
 F/2/Modified site: blocked amino end (Met) (in mature form) #status experimental
 F/108,265,289/Active site: Cys, His, Asn #status predicted

Query Match 69.5%; Score 2579.5; DB 1; Length 705;
 Best Local Similarity 66.6%; Pred. No. 1.8e-163;
 Matches 465; Conservative 111; Mismatches 121; Indels 1; Gaps 1;

QY 3 GIAAKLADREABAGLSHERAIKYNODYEALRNECLEAGTLFODPSFPAIPALGF 62
 DB 6 GIAAKLADREABAGLSHERAIKYNODYEALRNECLEAGTLFODPSFPAIPALGF 62
 QY 63 LEPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALGDCWLLAALASLTINNEI 122
 DB 63 LEPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALGDCWLLAALASLTINNEI 122
 QY 66 LEPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALGDCWLLAALASLTINNEI 125
 DB 66 LEPYSSKTRGMRKRPTEICADPOFTIGATRTDIOGALGDCWLLAALASLTINNEI 125
 QY 123 RYVPLNOSFOENYAGIFHFOFMQGEWVEVVDLPTKDGELLFVHSAEGSEFMSAL 182
 DB 123 RYVPLNOSFOENYAGIFHFOFMQGEWVEVVDLPTKDGELLFVHSAEGSEFMSAL 182
 QY 183 KAVAKINGCYEALSGGATTEGFEFTGGIAEWELKRPENLFKIIQALOKGSLG 242
 DB 183 KAVAKINGCYEALSGGATTEGFEFTGGIAEWELKRPENLFKIIQALOKGSLG 242
 QY 243 SIIITSAADSEAITPQKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 302
 DB 243 SIIITSAADSEAITPQKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 302
 QY 246 DITSAPDMEAVTFKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 305
 DB 246 DITSAPDMEAVTFKLVKGHAISVTGAEBVSNGSLQKILIRINPGEVEMTGRMND 305
 QY 303 SWNTIDPERERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDILTSDTYKKMKLTK 362
 DB 303 SWNTIDPERERELTRRHEDGEPFMSFDFLRHYSLEICNLTPDILTSDTYKKMKLTK 362
 QY 363 GNRGSGTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRRQR 422
 DB 363 GNRGSGTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRRQR 422
 QY 423 GEDMHTIGFAIYEVPEPFGQTNIHLKSNFELTNARAKSNFTFIMLREVLNRFKLPAG 482
 DB 423 GEDMHTIGFAIYEVPEPFGQTNIHLKSNFELTNARAKSNFTFIMLREVLNRFKLPAG 482
 QY 483 EYIVLSTPEPNDDGDCIRVPESEKADYQAVNDEIRANLEEPDISDDIDDOGRRLFAQ 541
 DB 483 EYIVLSTPEPNDDGDCIRVPESEKADYQAVNDEIRANLEEPDISDDIDDOGRRLFAQ 541
 QY 541 LAEDDAEISAFELQTIIRRLVLAQRDIDSGDSFISCTKIMWMDLSDSGSKLGKEFYIL 601
 DB 541 LAEDDAEISAFELQTIIRRLVLAQRDIDSGDSFISCTKIMWMDLSDSGSKLGKEFYIL 601

Db 546 AGEDMEISYFELKTIINRVIARHKD.IKTDFSLDCSRNWNVNLMDXGSGARLGLVERQITM 605

Qy 602 TKIQKYÖKIYREIDVDVRSGTMSNYEMRKALIEEAGFKMPCLQHGYIVARFADQLIIDFDN 661

Db 606 NKIRSWLTIFROYDDLDKSGTMSSEYEMRMALIESAGFKLNKLHQVVAVARADAETGVDFDN 665

Qy 662 FVRCVLRETLEPKLFQJDLBENTGETIELDLISLCSVSU 699

Db 666 FVCCLVKLETMFEEFFHSMDRGDTGVIAVMNLAELMLLLTM 703

RESULT 5

[illegible]

```

Db 1 LGPMSSKTYGIMKMRPFIBLSNPOFIYDGAIRTRTICGALGDCHULAIASLTINDTLTLH 132
Oy 123 RVPELINSFOENYAGIFHFQFQYGEWVEVVVDRLPTKOGELLFVHSABSGSEFMSALLE 182
Db 133 RVFPHGOSFOGNYAGIFHFQFQYGEWVDVVDLLPTKDGKLVFVHSABENEFMSALLE 192
Oy 183 KAYAKINGCYALAGGATTEGFEDPTGGIAEWYELKPKPPMLFKITQALOKGSLIGCSI 242
Db 193 KAYAKINGSYALAGGSGTSSEGFEDPTGVTWYELKRAPSDLYQIILKALRGSLIGCSI 252
Oy 243 DITSADSEAITFOKLKYGHAIVSYTGAEVEVNSGLQKILIRNPMGEVMTGRANDNCP 302
Db 253 DISSVLMEAITTFPKLVKGAHVSYTGAKOVYVRQVSVLIMRNPWGEVMTGAMSDSS 312
Oy 303 SMNTIDPEERELTRRHEDGEPFMSFSDFLHYSRLCINTLPTLTJSDTYKMKWLTAMD 362
Db 313 EMNVVDPEROQLKVKMGDEGFMSFRDPHMEFPLRELCTNLTPLAKSRITIRKNNTTLYE 372
Oy 363 GNMWRGSTAGGCRRYPNTFMNPOYLKLEBEDER--DGESGCTFLVGLIQKRRROR 420
Db 373 GTMRGSGTAGGCRRYPATFWNPOFKIRLDETDPPDYGDRESGGSFVLTALMOKRRER 432
Oy 421 KMGEDMHTIGIGIYEVPBELSGOTINHLSKNFPLTNRRESDPFINLREVLNPFKLPBG 480
Db 433 RFGMDMETIGIYAEVPELVGQPAVHLKRPFLPLANMSRANSEQFINLREVSTRFLRDPG 492
Oy 481 EYILVSTPEBNKGDGFCIRVSEKKADYQAVDDEIEANT--BEPDISHDIDDGVRRLPA 539
Db 493 EYVVVVPSTPEBNKGDGDFLRFPSSEKSGATVELDPOIQANLPDEBOVLSEERIDENFKALFR 552
Oy 540 QLAGEDAEISFPELOTILIRVLAARODIKSGPSLETCKIIVMDLMDSGSGKGLKGEFYI 599
Db 553 QLAGEDMEISYKELRTILNRILISKHODIRTGFSLESRSRSMVNLMDRNGNKLGIYENI 612
Oy 600 LMTKIQKQKITYREIDVDRSGTMSYEMRKALBEAGFMPQCLHQVIYARPADQOLITDF 659
Db 613 LMNRIIRNYLSIFRRFDLIDKSGMSGAYEMRWMAIESAGFLNKKLLEYLITRYSBDLAVDF 672
Oy 660 DNFRCLVRLTTLFKPIFKOLDPENTGTIELDLISML 695
Db 673 DNFVCCVLRLTMRFRFKITLDTDGDVVTFLPFKML 708

```

RESULT 6

calpain (BC:3.422.17) large chain 2, tissue-specific - rat
C|Species: Rattus norvegicus (Norway rat)
C|Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C|Accession: A48764; B48764
R|Sorimachi, H.; Ishiura, S.; Suzuki, K.
J. Biol. Chem. 268, 19476-19482, 1993
A|Title: A novel tissue-specific calpain species expressed predominantly in the stomach
A|Reference number: A48764, MUID:93374936, PMID:769035
A|Accession: A48764
A|Status: Preliminary
A|Molecule type: DNA; protein
A|Residues: 1-703 <SOR>
A|Cross-references: GB:D14479; NID:g441199; PIDN:BAA03370.1; PID:g441200
A|Experimental source: Ectomch
A|Note: Sequence extracted from NCBI backbone (NCBIN:137770, NCBIP:137771)
A|Accession: B48764
A|Status: Preliminary
A|Molecule type: DNA; protein
A|Residues: 1-379, 'SS', <SO2>
A|Cross-references: GB:D14480; NID:g441201; PIDN:BAA03371.1; PID:g495223
A|Experimental source: Ectomch
A|Note: Sequence extracted from NCBI backbone (NCBIN:137773, NCBIP:137775)
C|Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C|Keywords: alternative splicing; cysteine proteinase; EF hand; hydrolase
F:753-327/Domain: calpain catalytic domain homology <CAIP>
F:532-563/Domain: calmodulin repeat homology <EF1>
F:575-607/Domain: calmodulin repeat homology <EF2>

F:640-672/Domain: calmodulin repeat homology <EF3>
F:105,262,286/Active site: Cys, His, Asn #status predicted

Query Match	63.5%;	Score 2356.5;	DB 2;	Length 703;
Best Local Similarity	60.9%;	Pred. No. 1.2e-148;		
Matches 427;	Conservative 127;	Mismatches 140;	Indels 7;	Gaps 3;

[illegible]

RESULT 7

N:Altemate names: calpain chain 2 - rabbit (fragment)
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C:Date: 24-Jan-1988 #sequence_revision 24-Jan-1988 #text_change 16-Jul-1999
C:Accession: B24815
R:Emori, Y.; Kawasaki, H.; Sugihara, H.; Imajo, S.; Kawashima, S.; Suzuki, K.
J. Biol. Chem. 261, 9465-9471, 1986
A:Title: Isolation and sequence analyses of cDNA clones for the large subunits of two isoforms of rabbit calpain
A:Reference number: A92594; MUID:86250902; PMID:2424911
A:Accession: B24815
A:Molecule type: mRNA
A:Residues: 1-422 <EMO>
A:Cross-references: GB:ML13797; NID:g165665; P1DN:AAA31455.1; PID:g165666
C:Superfamily: calpain large chain calmodulin repeat homology; calpain catalytic domain
C:Keywords: calpain binding; cysteine proteinase; EF hand; hydrolase

F;251-282/Domain: calmodulin repeat homology <EF1>
F;294-326/Domain: calmodulin repeat homology <EF2>
F;327-356/Domain: calmodulin repeat homology <EF3>
F;359-391/Domain: calmodulin repeat homology <EF4>

Query Match	58.0%;	Score 2153;	DB 2;	Length 422;
Best Local Similarity	93.8%;	Pred. No. 1.9e-135;		
Matches 396;	Conservative 15;	Mismatches 11;	Indels 0;	Gaps 0;

QY	279	ÖKLIRINPMGEVEMTGRMNDNCPSMNTIDEEERERLTRRHDGSPMWSFSDPLRHYSL	338
Db	1	ÖKLIRINPMGEVEMTGRMNDNCPSMNTIDPEVERLAEKHEDGEPMWSFSDPLRHYSL	60
QY	339	EICNLTPDITLSDTYKKMKLTKMDGNMRGSGTACGRNYPNTFMWNPÖYILKEEBEDE	398
Db	61	EICNLTPDITLSDTYKKMKLTKMDGNMRGSGTACGRNYPNTFMWNPÖYILKEEBEDQ	120
QY	399	EDSGGCFVLGLÖKHRRRQRMKGEDMHTGPIYVPEBLSGOTNHLKSNFLLTRA	458
Db	121	EDSGGCFVLGLÖKHRRRQRMKGEDMHTGPIYVPEBLSQOTNHLKSNFLLTRA	180
QY	459	RERSDTFINLREVLNRFELPPGEYILVPSTEPMKGDPCIRVSEKKADYQADDEIEA	518
Db	181	RERSDTFINLREVLNRFELPPGEYILVPSTEPMKNGDFCIRVSEKKADYQADDEIEA	240
QY	519	NLEEFISDEDDIDGVRRLFAQLAGEBAEISAFELÖTLRRVLARÖDIKSDGFSIETCK	578
Db	241	NLEEFISDEDDIDGVRRLFAQLAGEBAEISAFELÖTLRRVLARÖDIKSDGFSIETCK	300
QY	579	IMVMMLSDSGSKLGLKEFYILMKIOYÖKYÖIYEBIDVRSGTMSYEMRALPEAGFKL	638
Db	301	IMVMMLSDSGSKLGLKEFYILMKIOYÖKYÖIYEBIDVRSGTMSYEMRALPEAGFKL	360
QY	639	PCÖLHOVIVAFADÖLIDFDNFCVRCIVLRETFKILFQÖDPENTGTIÖDLISWLCSFS	698
Db	361	PCÖLHOVIVAFADÖLIDFDNFCVRCIVLRETFKILFQÖDPENTGTIÖDLISWLCSFS	420
QY	699	VL 700	
Db	421	VL 422	

RESULT 6

calpain (EC 3.422.17) large chain 1 - chicken (fragment)
N.Alternate names: mu-calpain heavy chain
C.Species: Gallus gallus (chicken)
C.Date: 26-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C.Accession: S57195
R.Sotomachi, H.; Tsukahara, T.; Okada-Ban, M.; Sugita, H.; Ishiura, S.; Suzuki, K.
Biochim. Biophys. Acta 1261, 381-393, 1995
A.Title: Identification of a third ubiquitous calpain species - chicken muscle expressed
A.Reference number: S57194; MUID:95260862; PMID:7742367
A.Accession: S57195
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-586 <SOR>
A.Cross-references: EMBL:D38027; NID:S9882070; PTD:S9882071
C.Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
C.Keywords: cysteine proteinase; EF hand; hydrolyase
F.1-508/Domain: calpain catalytic repeat homology (fragment) <CALP>
F.457-489/Domain: calmodulin repeat homology <EF1>

Query Match	53.4%;	Score 1984;	DB 2;	Length 586;
-------------	--------	-------------	-------	-------------

Matches 354; Conservative 108; Mismatches 114; Indels 4; Gaps 2;

QY 120 LTRVVPENQSFQENYAGIIFHHQFQOYIBWMAVVDLDFTRDELLFVSHABSEFMSA 179
Db 1 LTRRVVPHQGSFQNGIAGIFHHQIWOPEBMQDVVDYDLPTRKDGKLLFVSHABSEFMSA 60
QY 180 LTRKAVKINGCEYALSGATTEGPEDFTGGAENYELKKRPPVLFKIQALQKSLG 239

Db 61 LLEKAVKAVNGCYEALSGSTSEGFEDFTGVTWYDLRKRPADLYOINIKALERSGLG 120
 Qy 240 CSIDITSAADSEALTFOKLVKGHAYSVTGAEEVENSGLQKILIRNPMGEVETGRMD 299
 Db 121 CSIDITSAFDMKATFKKMKVGHAYSVTGAQOISYRQOSIGLIMKRPMEGEVETGRMD 180
 Qy 300 NCPSMNTIDPEERRLTRRHEDGEPWMSFSDFLRHYSRLICNLTPDTLTSITYKKWLT 359
 Db 181 SSSQWNEVEPSLRQIWMEDGEPWMSFADFLNEFRLRICILTPALQSRKFRKMNTR 240
 Qy 360 KMDGNMRGSGTAGCCRRYPTFMNPOYLKLEDEDEDEP--GEGCTFLVLIOKRR 416
 Db 241 LYDSMWRGSGTAGCCRRYPTFMNPOYKCLEEVDDGDPGREGCCSFLALMOKHR 300
 Qy 417 RROKMGEDMHTTIFGVEVEEELSGQTNHLSKNFELTNARERSDFINLRVLRPFK 476
 Db 301 RRRRYGKDMETTFPANYVEVPEVHGRSGVHLQDFELSNASRRASEOFILNRVSTRLR 360
 Qy 477 LPPEGYILVPSFEPNKGDFCIRVFSEKKADYQAVDELEANL--EEFDISEDDIDDGVR 535
 Db 361 LPPEGYILVPSFEPNKGDFCIRVFSEKKADYQAVDELEANL--EEFDISEDDIDDGVR 535
 Qy 536 RLPLQAGDAEIAFELQTLIRVLAQRDIXSDGSIETCKIMVMDLSDSGSKGLK 595
 Db 421 ALFRLQAGDPMELISVTELTQTLIRVLAQRDIXSDGSIETCKIMVMDLSDSGSKGLK 595
 Qy 596 EFLYLMTRIKOKYREIDVRSGTMSYEMRKALEAGFMPCOLHOVTVARFADQL 655
 Db 481 EFLYLMTRIKOKYREIDVRSGTMSYEMRKALEAGFMPCOLHOVTVARFADQL 655
 Qy 656 IIPDNFVRCLVRLTLEFKIFKQDPENTGTIELDLISWL 695
 Db 541 AIDFDSFVCCVRLTLEFKIFKQDPENTGTIELDLISWL 695

RESULT 9

S57196
 calpain (EC 3.4.22.17) large chain 3 - chicken
 N:Alternate names: calpain p94 heavy chain; n-calpain-1 heavy chain
 C:Species: Gallus gallus (chicken)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
 C/Accession: S57196
 R:Scrinach, H.; Tsukihara, T.; Okada-Ban, M.; Sugita, H.; Ishiura, S.; Suzuki, K.
 Biochim. Biophys. Acta 1261, 381-393, 1995
 A:Title: Identification of a third ubiquitous calpain species - chicken muscle expressed
 A:Reference number: S57194; MUID:95260862; PMID:7742367
 A/Accession: S57196
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-810 <SOR>
 A/Cross-references: EMBL:D30828; NID:G882072; PIDN:BA07230.1; PID:G1552167
 C:Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C:Keywords: cysteine protease; EF hand; hydrolase
 F:93-393/Domain: calpain catalytic domain homology <CALP>
 F:746-778/Domain: calmodulin repeat homology <EFH>

Query Match 51.1%; Score 1897; DB 1; Length 810;
 Best Local Similarity 45.5%; Pred. No. 4.8e-118;
 Matches 350; Conservative 145; Mismatches 177; Indels 98; Gaps 7;

Qy 18 LGSHERAIKYLNDYVLEALRNECLAGTLFQDPSPFAPISALGFYELGYSKTRGMWRK 77
 Db 41 ISRNQPIIKVKEKTEELHKKCLFENILYEDPDPNPTSIFYGQKVPK-----FEKRR 95
 Qy 78 PTEICADPQITIGATRTDIOGALGDCWLLAIALSLTNEELIARVPLNQSFOENYAG 137
 Db 96 PREICEPPIIGANRTDIOGELGDCWFLAIALSLTNEELIARVPLNQSFOENYAG 155
 Qy 138 IFPHQPOYGEWVAVDDRLPTKQGLLFVHBSGSEFMSALKEKVAKINGCYEALSG 197
 Db 156 IFPHQPOYGEWVAVDDRLPTKQGLLFVHBSGSEFMSALKEKVAKINGCYEALSG 215
 Qy 198 GATTEGEDFTGIAEWELKKPPNLFKTIQALQKSLGCSID----- 243

Db 216 GNTTEAMEDFTGVIIFYEIKDAPKDIYKIMKIALRGSLMSSIDNLFHYGAAPRSD 275
 Qy 244 -----LISAADSEALTFOK-----LVKGHAYSVTGA 270
 Db 276 IGEILIRNKNVKNLNAQMTTISTYDYGTDPERPANTIMPQYETRMSCGLVGHAYSVTAYE 335
 Qy 271 EVENSGLQKILIRNPMGEVETGRMDNCPSEMTIDPEERRLTR--HEDGEPWMSFS 329
 Db 336 ETTKGEKORVLRNPMGEVETGRMDNCPSEMTIDPEERRLTR--HEDGEPWMSFS 329
 Qy 330 DFLRHYSRLICNLTPDTLTSITYKKWLTAKDGNRRSGTAGCCRRYPTFMNPOYL 389
 Db 396 DFLRHYSRLICNLTPDTLTSITYKKWLTAKDGNRRSGTAGCCRRYPTFMNPOYL 389
 Qy 390 KLEEDDEDEDESGCTPLVGLIOKRRROROKMGEMHTTIGFIYVEPEELSGQTNHLS 449
 Db 456 KLEEDDEDEDESGCTPLVGLIOKRRROROKMGEMHTTIGFIYVEPEELSGQTNHLS 449
 Qy 450 KNEFLTNARERSDFINLRVLRPFKLPPEGYILVPSFEPNKGDFCIRVFSEKKADY 509
 Db 515 KDFFLYNASARKSTIYNNREISERFLPPEGYILVPSFEPNKGDFCIRVFSEKKADY 574
 Qy 510 QAVDELEANL-----EEFDISEDDIDDG----- 533
 Db 575 EEVENNIEADRPSEKKKGRPIIFVSDRANSNKLTTDEBAGKGEKTHVDEKKRSSAKAR 634
 Qy 534 -----VRLFAOLAGDAEISAFELQTLIRVLAQRDIXSDGSIETCKIMVMDL 585
 Db 635 EKSEETOPFNIRQLAGDMEICREELRVNLNVKAKHDLTBGELEBSRSMALMD 694
 Qy 586 SDGSKGLKEFYILWTKIOKYREIDVRSGTMSYEMRKALEAGFMPCOLHOVTVARFADQL 655
 Db 695 TDSGKINPDEFRLMDKIKSWQKIFKHYADHSGTINSTEMNNAVKDAGFRNLNQLDYD 754
 Qy 646 IVARFADQLIIPDNFVRCLVRLTLEFKIFKQDPENTGTIELDLISWL 695
 Db 755 ITWRVADKNNIDPDSFICCFVRLDAMFRAFHADKDXGDIKLVLEWL 804

RESULT 10

B34488
 calpain (EC 3.4.22.17) large chain 3 - rat
 N:Alternate names: Rattus norvegicus (Norway rat)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: B34488; S10589
 R:Scrinach, H.; Imajoh-Ohmi, S.; Emori, Y.; Kawasaki, H.; Ohno, S.; Minami, Y.; Suzuki, J.
 Biol. Chem. 264, 20106-20111, 1989
 A:Title: Molecular cloning of a novel mammalian calcium-dependent protease distinct from
 A:Reference number: A94688; MUID:90062125; PMID:2555341
 A/Accession: B34488
 A:Molecule type: mRNA
 A:Residues: 1-821 <SOR>
 A/Cross-references: GB:005121; NID:G205955; PIDN:AAA1790.1; PID:G205956
 R:Scrinach, H.; Ohmi, S.; Emori, Y.; Kawasaki, H.; Saido, T.C.; Ohno, S.; Minami, Y.; Suzuki, J.
 Biol. Chem. Hoppe-Seyler 371(Suppl.), 171-176, 1990
 A:Title: A novel member of the calcium-dependent cysteine protease family.
 A:Reference number: S10589; MUID:90380278; PMID:2400579
 A/Accession: S10589
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-50, 'E', 52-211, 'V', 213-252, 'K', 254-821 <SOR>
 C:Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C:Keywords: calcium binding; cysteine protease; EF hand; hydrolase
 F:99-400/Domain: calpain catalytic domain homology #status atypical <CALP>
 F:649-680/Domain: calmodulin repeat homology <EF1>
 F:692-724/Domain: calmodulin repeat homology <EF2>
 F:725-754/Domain: calmodulin repeat homology <EF3>
 F:757-789/Domain: calmodulin repeat homology <EF4>
 F:790-821/Domain: calmodulin repeat homology <EF5>
 F:129,334,358/Active site: Cys, His, Asn #status predicted

Db 654 FRNFKOIAAGDMIEICADELKKVLTNVNKKHDKLTHGFTLSECRSMIALMDTGGSKLN 713
 QY 594 LKEEYLIAMTKIOKQKTYREIDVDKSGMNSYEMRKALBEGAFKPMQOLHOVIARPAD 653
 Db 714 LQEHHLNNKIKAMQKFKHYDIDQSGTINSYERNANVNDGFHNNOLYDIITWYADK 773
 QY 654 QLIIDPNFVACLVRLETFKIFKQDPENTGTELDLISWL 695
 Db 774 HNNIDFDFSCFVRLBGMFRAFIAPFDKDGDIILKLVLEWL 815

RESULT 12

A55054
 calpain (EC 3.4.22.17) large chain - fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
 C/Accession: A55054
 R/Emori, Y.; Saito, K.
 J. Biol. Chem. 269, 25137-25142, 1994
 A/Title: Calpain localization changes in coordination with actin-related cytoskeletal ch
 A/Reference number: A55054; MUID:95014293; PMID:7929201
 A/Accession: A55054
 A/Status: preliminary; nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-805 <EMO>
 A/Cross-references: GB:X78555; NID:g562287; PIDN:CAAS5297.1; PID:g562288
 C/Genetics:
 A/Gene: FlyBase:Calpa
 A/Cross-references: FlyBase:FBgn0012051
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: cysteine proteinase; duplication; EF hand; hydrolyase
 F/90-347/Domain: calpain catalytic domain homology <CALP>
 F/676-708/Domain: calmodulin repeat homology <EF1>
 F/709-738/Domain: calmodulin repeat homology <EF2>
 F/741-773/Domain: calmodulin repeat homology <EF3>
 F/774-805/Domain: calmodulin repeat homology <EF4>

Query Match 42.9%; Score 1594.5; DB 1; Length 805;
 Best Local Similarity 41.2%; Pred. No. 6e-98;
 Matches 322; Conservative 133; Mismatches 209; Indels 117; Gaps 14;

QY 13 EAAGLGSHERAIVKYNODYALNNECLAEGLTFQDPSFPAIPSAFGKELGPKYSSKTRG 72
 Db 38 EKSSSLGPKPYSEV-----QDYETIINSCLASGLFEDPLFPASNSLQCSR-----RPRRH 87
 QY 73 MRWRPPEICADPQPIIGATRTDICGALGDCMLLAIASTLTNEBILAVPLNLSQ 132
 Db 88 IEWLRPHIAPNPQFVEGYSRFPVQOGLDCMLLATANLTQESNLFERVIAPEQSF 147
 QY 133 ENVAGIFHFQFQWGEWVVDRLPTKQDELLFVHSAEGSEFWSALLERKAVAKINGCY 192
 Db 148 ENVAGIFHFQFQWGEWVVDRLPTKQDELLFVHSAEGSEFWSALLERKAVAKINGCY 207
 QY 193 EALSAGATGEGEDFTGIAEWELKRPPLFKIIQKALOKSGILGCSIDITSAADSEA 252
 Db 208 EALGSGSTCEMEBFTGSEWYDLKEAPNLFTILQKAARNMMSGCSIE-PDPNVEA 266
 QY 253 ITPKLVKGAHYSTVGAEEVE-----SNGLQKILIRPNPG-EVEWNGRNNDCPSMNT 306
 Db 267 ETPOGLRGHAYSTVTKCLDITVTPNRQKI-PMIRKMPGNAGNAEWNGPWSDSPEKRY 325
 QY 307 IDPERERLTRLRHE-DGEFWMSPFDPLRHYRLKLEICNLTPTLTSDTY----KKMKLTLM 361
 Db 326 IPEQKXAEIGLTPRDRGEFWMSPFDPLRHYRLKLEICNLTPTLTSDTY----KKMKLTLM 385
 QY 362 DGNRRSGTAGGCNYPVTFMNFQYLIKLEEDDEDEDESGCTFLVGLIIOKRRQK 421
 Db 386 EGEWTPVTAAGCCNFPDLTFMNFQYLIKLEEDDEDEDESGCTFLVGLIIOKRRQK 443
 QY 422 MGEWMHTIGRGIYVPE-ELSGQINIHLSKNFPLTNARRESDFINIRETLNFKLP 480
 Db 444 MGEWCLTIGAIYSLNDRELENRPO--GLNFFRYKSSVGRSPHINTREVCAKFKLP 500

RESULT 13

A39343
 calpain (EC 3.4.22.17) large chain - fluke (Schistosoma mansoni)
 C/Species: Schistosoma mansoni
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C/Accession: A39343; A45642
 R/Andersen, K.; Tom, T.D.; Strand, M.
 J. Biol. Chem. 266, 15085-15090, 1991
 A/Title: Characterization of cDNA clones encoding a novel calcium-activated neutral prot
 A/Reference number: A39343; MUID:91332027; PMID:1669543
 A/Accession: A39343
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-758 <AND>
 A/Cross-references: GB:M67499; NID:g160936; PIDN:AAA29858.1; PID:g160937
 R/Karcz, S.R.; Podesta, R.B.; Siddiqui, A.A.; Dekaban, G.A.; Strejan, G.H.; Clarke, M.W
 Mol. Biochem. Parasitol. 49, 333-336, 1991
 A/Title: Molecular cloning and sequence analysis of a calcium-activated neutral protease
 A/Reference number: A45642; MUID:92131071; PMID:1775175
 A/Accession: A45642
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-14, 'A', 16-120, 'A', 121-127, 'H', 129-213, 'R', 215-384, 'VTC', 388-440, 'S', 442-758
 A/Cross-references: GB:M74233; NID:g160934
 A/Note: sequence extracted from NCBI backbone (NCBI:79194)
 C/Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain
 C/Keywords: calcium binding; cysteine proteinase; duplication; EF hand; hydrolyase
 F/14-380/Domain: calpain catalytic domain homology <CALP>
 F/628-660/Domain: calmodulin repeat homology <EF1>
 F/661-690/Domain: calmodulin repeat homology <EF2>
 F/693-726/Domain: calmodulin repeat homology <EF3>
 F/727-758/Domain: calmodulin repeat homology <EF4>
 F/14,313,337/Active site: Cys, His, Asn #status predicted

Query Match 32.3%; Score 1199.5; DB 1; Length 758;
 Best Local Similarity 38.6%; Pred. No. 9.5e-72;
 Matches 279; Conservative 136; Mismatches 234; Indels 73; Gaps 26;

QY 16 EAGLSHERAKIYNODYEALNNECLAEGLTFQDPSFPAIPSAFG-----FKELOPYSKTR 71
 Db 70 KQPAKFLMNVNVAQKQETIVKRLKTERLTWEDPDPADKALGNLPDFE----- 120
 QY 72 GMEWKPTEICADPQPIIGATRTDICGALGDCMLLAIASTLTNEBILAVPLNLSQ 131
 Db 121 RIEMKRPLEINPAKFPAGASRFDIEGALGDCMLLAIVASISGTFQFLDYVPPQDEL 180
 QY 132 Q-ENVAGIFHFQFQWGEWVVDRLPTKQD-ELLFVHSAEGSEFWSALLERKAVAKI 188

Db 181 KGEYVGVVRRFRFGHWVEVLIDRLPVROGINTLVFMSNDPTFWSALKEKAYAKL 240

QY 189 NGCEYALSGATTEGFEPTGGIAEMWYL--KKPPNLFKIIQKALQKSLGCSIDITS 246

Db 241 NGCYAHLSSGSGSASAMDLTGICLSLELNKERRPSLDIOLKTYARCCIMGCSID--- 297

QY 247 AADSEAITFOK---LYKHAIVSYTGAEVSSNSLOKLIRINPMGEV-EMTGRMNDNC 301

Db 298 ----SSVMEQKMDNGLIGSHAYSLTGYPVAVYRGRITWMLRNPWMDSHKMGAMDGS 353

QY 302 PSNMTIDPEERERLTRRH-EDGEFMSFSDLRHYSLEICNLTPDTL--TSQYKKKNKL 358

Db 354 POREISEQEKKNINLSPTADGEFMSYEDFCYFSRVEVCHGLGLESLEYNQNHGRRL 413

QY 359 TK--MDGNMRGSTRAGCCRYNPNTFMNPOYLKLEEBEDEDESGCTFLVGLIQHR 416

Db 414 DEAFISGQGRNVAVAGCINNRITTYNTNPOFRITVEDPDDDNK--CSVLIGLMQTDI 471

QY 417 RRQRKMGEDMTTIGFIYEVPEELSGQTNHLSKNPFLNRRARSPTFINLRVLRPK 476

Db 472 R--KKVADGADPOPIGFMYNAPDDL---NTLSRAQLLTSPLAKSQ--FINTREVTQAFR 524

QY 477 LPGEYILVSTPEPNKDGFCIRVFESEKADYAV-----DDEIEA-NLEEFDI 525

Db 525 VPEGSYVIPSITPDENIEVNFILKVFQTSITTEQLEEDNTNQGLPDVILEALKEDTTL 584

QY 526 SEDDIDGVRRLFAQLAGEDAE--ISAPEIQTILRRVLAKRQDLS--DGSFLETCKIMV 581

Db 585 DE---DEIEQKF--LAIRPKTNAINAVKLGELINN--SLQDIPRFGFNEKLSMNV 637

QY 582 DMLDSGSGKLGKEPIITMTKIQKQKTYREIDVDSGTMSYEMRKALBEAGFKMPCQ 641

Db 638 ASVDNNLTGVHEINFEFMDLWIOAKGMGHIPIKHEVDQSGYFAVEFREALNDAGYHSNR 697

QY 642 LHQIVARFAD---DQILIDPNVRCVLETLTKFKQKDPEN-IGTIEBLDISLWLCF 697

Db 698 LINAIRKYDPGTDK--ISFEDPMLCMVRKTAFTI-EAHPKNIEGTSLFSAEDYLRF 754

QY 698 SV 699

Db 755 RV 756

RESULT 14

C06G4.2 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C/Accession: S44749

R:Waterston, R.

Submitted to the EMBL Data Library, November 1993

A:Description: Sequence of the C. elegans cosmid C06G4.

A:Reference number: S44747

A:Accession: S44749

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-653 <WAT>

A:Cross-references: EMBL:L25598; NID:G409290; PIDN:AAA27940.1; PID:G409293

C:Genetics:

A:Insertions: 26/1; 69/1; 216/3; 269/3; 357/2; 401/2; 558/1; 583/3; 613/3

C:Superfamily: C06G4.2 protein; calpain catalytic domain homology

F:245-498/Domain: calpain catalytic domain homology <CALP>

Query Match 31.9%; Score 1184; DB 1; Length 653;

Best Local Similarity 47.3%; Pred. No. 8.3e-71;

Matches 231; Conservative 73; Mismatches 124; Indels 60; Gaps 9;

QY 30 QDYEALENECTLEAGTLFQDPSFPAIPALGKELGYPSSKTRGMRMRKPTIEICADPOFI 89

Db 205 QDFEHLADQGLSKRLPFDPOFLANDSSLFSSKRP-----KRYEMLRPGEITREPDLIT 259

QY 90 GGATRTDICGALGDCMLAALASLTNNEETLARVPLNOSPOENVAGIFHPQFQYGEW 149

Db 260 EGHSRDVIYQGLGDCMLAAANLTLKDELFYRVVPPDGSFTENYAGIFHPQFQYGEW 319

QY 150 VEYVVDRLPTKXGELLFVHASAGSEFWASALKEAKVINGCYEALSGATTEGFEPTG 209

Db 320 VDVIIDRLPTNSGELLVHMASNNEFWASALKEAKVINGCYEALSGATTEGFEPTG 379

QY 210 GIAEWELKKPPNLFKIIQKALQKSLGCSIDITSADS---EAITFOKLYGHAYSV 266

Db 380 GLTEFIDLKPPNLMQMMWRGFMGSLFCGSIIE---ADPNWEAMSGVLKGYHAYSI 435

QY 267 TGAEVSSNSLOKLIRINPMG-EVMTGRMNDNCPSNMTI--DPERERLTRRHEDEGEF 324

Db 436 TGRIVDPNGQTCILIRINPMGNEQWMPMSDNRKMSVDSYKQDGKLFDDHGEF 495

QY 325 WMSFDFLRHYSRLICNLTPDTLTSDTYK-----WKTQKDNRRSGTAGCCR 375

Db 496 WMSFDFMRFRFECMEICNLGPDYM-DEVYQMTGVKAAQVMAANTHDGAMVRQTAGCCR 554

QY 376 NYPTFMNPOYLKLEEBEDEDESGCTFLVGLIQHRRRQRKMGEDMTIGFIYE 435

Db 555 NYITFANNPQFRVQLTSDPDDD--ELCT----- 583

QY 436 VPEELSGQTNHLSKNPFLNRRARSPTFINLRVLRPKLPGEYILVSTPEPNKDG 495

Db 584 ----AGNNRGRLSKQFPAKNSAMRSAAFINLRMTGRVRPQVYVVPSTPEPNEA 638

QY 496 DFCIRVFS 503

Db 639 EFMRLRYT 646

RESULT 15

A24815

calpain (BC 3.4.22.17) large chain 1 - rabbit (fragments)

N:Alternate names: calcium-activated neutral proteinase (CANP)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 24-Jan-1988 #sequence_revision 24-Feb-1994 #text_change 16-Jul-1999

C/Accession: A24815; A41418

R:Emori, Y.; Kawasaki, H.; Sugihara, H.; Imajoh, S.; Kawashima, S.; Suzuki, K.

J. Biol. Chem. 261, 9465-9471, 1986

A:Title: Isolation and sequence analyses of cDNA clones for the large subunits of two 1

A:Reference number: A92594; WUID:86250902; PMID:2424911

A:Molecule type: mRNA

A:Accession: A24815

A:Residues: 19-320 <EMO>

A:Cross-references: GB:M13363; NID:G165667; PIDN:AAA1456.1; PID:G165668

R:Kawasaki, H.; Imajoh, S.; Suzuki, K.

J. Biochem. 102, 393-400, 1987

A:Title: Separation of peptides on the basis of the difference in positive charge: simul

A:Reference number: A41418; WUID:88032960; PMID:3667575

A:Accession: A41418

A:Molecule type: protein

A:Residues: 1-18;125-154;313-320 <KAW>

A>Note: Sequence was deduced from composition by homology

C:Superfamily: calpain large chain; calmodulin repeat homology; calpain catalytic domain

C:Keywords: blocked amino end; calcium binding; cysteine proteinase; duplication; EF hai

F:148-179/Domain: calmodulin repeat homology <EF1>

F:191-223/Domain: calmodulin repeat homology <EF2>

F:224-253/Domain: calmodulin repeat homology <EF3>

F:256-288/Domain: calmodulin repeat homology <EF4>

F:289-320/Domain: calmodulin repeat homology <EF5>

Query Match 24.9%; Score 925.5; DB 2; Length 320;

Best Local Similarity 56.6%; Pred. No. 4.5e-54;

Matches 167; Conservative 64; Mismatches 63; Indels 1; Gaps 1;

QY 402 ESGCTFLVGLIQHRRRQRKMGEDMTTIGFIYEVPEELSGQTNHLSKNPFLTNRRAR 461

Db 20 ESGCSYVALMQGHRRRRRFRGDMETIGAVYEVRELVGQALMLKDKDFLANSRRAR 79

QY 462 SDTFINLRVLRNFKLPGEYILVSTPEPNKDGFCIRVFSKKADYQAVDDEIENL- 520

Db 80 SEQFINLRVSTFRPLPGEYVVPSTPEPNKGDVLRFFSKRAGTQELDQIOANLP 139

```

QY 521 EEPDISEDDIDDGVRLFAQLAGEDEISAFELQTIIRVLAKRODIKSGFSIETCKIM 580
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 DEQVLSAEIENFKALFRQLAGEDLEISVRELQTIILNRITSKHDLRTKGFMSQCRSM 199
QY 581 VDMIDSDGSGKLGKEFYIIMTKIOKYOKIYREIDVDRSGTMSYEMRKALEEAGFKQPC 640
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 VNLMDRDGNGKLGVEFNILMNRIRNYLAIFRKFDLDKSGSMSAYEMRMAIESAGFKLTK 259
QY 641 QLHOVIARFADDDLIIDFNFVRCLVRLFTLFIKIPQOLDPENTGTIELDLISWL 695
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 KLYELIITRSEPDLAVDNFCCLVRLFTLFIKIPQOLDPENTGTIELDLISWL 314

```

Search completed: July 24, 2003, 12:53:49
 Job time : 37.3469 secs

APPLICANT: Hitachi, Ltd.
FILE OF INVENTION: AIP1-INTERACTING PROTEINS AND USE THEREOF
FILE REFERENCE: HITA.0164
CURRENT APPLICATION NUMBER: US/10/370,481
PRIOR FILING DATE: 2003-02-24
PRIOR APPLICATION NUMBER: US 60/358,369
PRIOR FILING DATE: 2002-02-22
PRIOR APPLICATION NUMBER: US 60/019,808
PRIOR FILING DATE: 2002-09-03
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 21
LENGTH: 191
TYPE: PRT
ORGANISM: Homo sapiens
US-10-370-481-21

Query Match 4.5%; Score 168; DB 6; Length 191;
Best Local Similarity 25.0%; Pred. No. 2.7e-08;
Matches 41; Conservative 39; Mismatches 68; Indels 16; Gaps 3;

QY 553 LQTLRLRYLAKRQDIKSD-----GFSIETCKIMVMDLSDSGSKGLKERYI 599
DB 28 LNVFQKRVKDSGVISDTLQALSNGTPTFNPVTVRSITISMPDRNKAGVNFSEFTG 87
QY 600 LMTKIOKQKYREIDVDKSGTMSYEMRKALBEAGFPMPCQLHQVIVARF-ADDQLIID 658
DB 88 VMKVTDMQVNFRTYDRNSGMIDKELKQALSGGVRLSDQPHDILIRKDRGRCGLA 147
QY 659 FDNVRCVLRLETFLKIFKQIDPENTGTIELDISWL--CNSVL 700
DB 148 FDDIIGGCIYLRLLDIFRRYDTQDQGWIVQSYEYLSMVSIV 191

RESULT 3
US-60-479-073-498
Sequence 498; Application US/60479073
GENERAL INFORMATION:
APPLICANT: De Wilde, Gert Jules Hector
APPLICANT: Saunders, Michael John Scott
TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
FILE REFERENCE: D00590.70042.US
CURRENT APPLICATION NUMBER: US/60/479,073
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 526
SOFTWARE: PatentIn version 3.2
SEQ ID NO 498
LENGTH: 5373
TYPE: PRT
ORGANISM: Homo sapiens
US-60-479-073-498

Query Match 3.2%; Score 117; DB 7; Length 5373;
Best Local Similarity 19.5%; Pred. No. 0.28;
Matches 114; Conservative 84; Mismatches 214; Indels 172; Gaps 24;

QY 180 LLEKAYAKINGCYEALSGGATTEGFEFTGSIAMWELKPPNLFKTIQALQKSLG 239
DB 2659 LEDLAADRINRLQAL--ASTQFOQMFDLRTWLDK-----QSQAKNCPIIS 2705
QY 240 CSIDITSADSEAITFOKLKVGH--AVSYTGAE-----EVESNGSLQ-KLIRIRNP 287
DB 2706 AKERLRSQLOENBEFQKSLNQHSGSYEIVVAEGBSLLISVPGEKRTQLONLVELKNH 2765
QY 288 WGE-----VEWTGRWNCPSMN-----TIDPERERLRLTR 317
DB 2766 WEELSKTADRSQRLKDCMQKQKQYQWNVHEDLVPMIBCKAKMELARTLDPVQLSSLL 2825
QY 318 RHEDGEFWMSPDLRHSRLICNLTPDTLTSPTYKKMKLTKMDGNWRGSIAGGCRNY 377

DB 2826 RSK-----AMLNEVERKRSLLLEILNSAAD----- 2849
QY 378 PNTFWMNPQYLILKEEDEDDEDESGCTFLVGLIQHRRRQRKMGEDMTTIGFYIEVP 437
DB 2850 -----ILINSSADED-----G-----RDKAG-----INQNDAYT 2877
QY 438 EELSGQTNHLSKNFLTNARERSDTFINLRVLRNFKLPGEYILVSTPEPNKGDG 497
DB 2878 EELQAKTG-SLEE--MTQLRFQESFKNIKKVGAQKQLRIFALGSAQCSNNK--- 2930
QY 498 CIRVSEKQADYQAVDEITANIEEF-DISEDDIDGVRFLPQALAGEDEIFAELQTI 556
DB 2931 -----LEKRAQGEVQALEPPQVYLRNFTQGLVEADPDGSDSLIHQAEVAQOEFLV 2985
QY 557 LRRVLA--KRQDIKSDGFSIETCKI-----MYDMLDS-DGSGKLQ----- 593
DB 2986 KQVNSGCVMMENKLEBIGFHCVRVEMFSQLDLDELDMGALIRGDRDSLOSQIEDVR 3045
QY 594 --LKEFYILMTKIOKQKYREIDVDKSGTMSYEMRKALBEAGFPMPCQLHQVIVARFA 651
DB 3046 LFLNKIHLKLDIEASBAECRHW-LBEGTLDLGLKRELEALN-KQCGKLTGRGKR-- 3101
QY 652 DDQLIIDFQNFVRCVLRLETFLKIFKQIDPENTGTIELDISWL 695
DB 3102 QEOLEL-----TLGRVEDPYRKLKGLNDATTAAEAEALQWV 3138

RESULT 4
US-60-479-073-500
Sequence 500; Application US/60479073
GENERAL INFORMATION:
APPLICANT: De Wilde, Gert Jules Hector
APPLICANT: Saunders, Michael John Scott
TITLE OF INVENTION: Amino acid sequences useful for developing compounds for the
TITLE OF INVENTION: prevention and/or treatment of metabolic diseases and nucleotide
FILE REFERENCE: D00590.70042.US
CURRENT APPLICATION NUMBER: US/60/479,073
CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 526
SOFTWARE: PatentIn version 3.2
SEQ ID NO 500
LENGTH: 5430
TYPE: PRT
ORGANISM: Homo sapiens
US-60-479-073-500

Query Match 3.2%; Score 117; DB 7; Length 5430;
Best Local Similarity 19.5%; Pred. No. 0.28;
Matches 114; Conservative 84; Mismatches 214; Indels 172; Gaps 24;

QY 180 LLEKAYAKINGCYEALSGGATTEGFEFTGSIAMWELKPPNLFKTIQALQKSLG 239
DB 2715 LEDLAADRINRLQAL--ASTQFOQMFDLRTWLDK-----QSQAKNCPIIS 2761
QY 240 CSIDITSADSEAITFOKLKVGH--AVSYTGAE-----EVESNGSLQ-KLIRIRNP 287
DB 2762 AKERLRSQLOENBEFQKSLNQHSGSYEIVVAEGBSLLISVPGEKRTQLONLVELKNH 2821
QY 288 WGE-----VEWTGRWNCPSMN-----TIDPERERLRLTR 317
DB 2822 WEELSKTADRSQRLKDCMQKQKQYQWNVHEDLVPMIEDCKAKMSELRTLDPVQLSSLL 2881
QY 318 RHEDGEFWMSPDLRHSRLICNLTPDTLTSPTYKKMKLTKMDGNWRGSIAGGCRNY 377
DB 2882 RSK-----AMLNEVERKRSLLLEILNSAAD----- 2905
QY 378 PNTFWMNPQYLILKEEDEDDEDESGCTFLVGLIQHRRRQRKMGEDMTTIGFYIEVP 437
DB 2906 -----ILINSSADED-----G-----RDKAG-----INQNDAYT 2933
QY 438 EELSGQTNHLSKNFLTNARERSDTFINLRVLRNFKLPGEYILVSTPEPNKGDG 497

```

Db      2934 BELQAKTG-SLEB-----WTQRLREFQESFKNIKKVEGAKQLEIFDALGQACSNKN--- 2986
Qy      498 CIRVSEKKAQYQAVDEIEANLEEF-DISEDDIDGVRRLFAQLAGEDAISAFELQTI 556
Db      2987 -----LEKRAQGEVLQALBPQVDYLRNFTQGLVEDAPDSQSDASQLHQLEVAQOEFLV 3041
Qy      557 LRRVLA-KKQDIKSDGFSIETCKI-----MVMMLDS-DGSGKLQ----- 593
Db      3042 KQVNSGCVAMENKLEGIQGFHCVRFPMSQLADLDELQMGALGRDTSLSQSIEDVR 3101
Qy      594 --LKEFIIITKIQYQKIRREIDVDSGTMSYEMKALBEAGFKPCCQLHOVYARFA 651
Db      3102 LFNKHLHVLKDLIEASEACRHM-LBEEGTLDLGLKRELEALN-KQCGKLTGRKAR-- 3157
Qy      652 DDOLIIDFNFVACLVLETLFKIFKQJDEPENTGTIELDLISWL 695
Db      3158 QEOLEL-----TLGRVEDFYRKLGKLNATTAABEAALQNV 3194

```

```

RESULT 5
US-10-273-573-8670
; Sequence 8670, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ. ID NOS: 10994
; SOFTWARE: Custom
; SEQ. ID NO: 8670
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(259)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-8670

```

```

Query Match      2.9%; Score 108.5; DB 6; Length 259;
Best Local Similarity 28.4%; Pred. No. 0.013;
Matches 42; Conservative 24; Mismatches 61; Indels 21; Gaps 7;
Qy      538 FAQAGEDAEISAFELQTIILRVLAQRQDIKSDG---FSIETCKIWMVMDLSDSGKLG 594
Db      61 FAVAGQDQIDDELQRC-----TQSGIAGGYKPFNLETCLRWVSMIDRMSGTMR 114
Qy      595 K-EFYL-----WTQIKYQKIRREIDVDSGTMSYEMKALBEAGF-KMPCCQLHOV 645
Db      115 SIEFKLLGLLEWAGROHIFSTDRWEQXDPQELP--GPXQTMGFXYVPCGXIS 172
Qy      646 IVARFADQIIDFNFVACLVLETLF 673
Db      173 FAKRYSTNGK-ITFDYIACCVQTXGVF 199

```

```

RESULT 6
US-10-280-962A-45
; Sequence 45, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorige, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2342
; CURRENT APPLICATION NUMBER: US/10/280,962A
; CURRENT FILING DATE: 2002-10-25
; NUMBER OF SEQ. ID NOS: 56
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ. ID NO: 45
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Thermococcus sp.
US-10-280-962A-45

```

```

Query Match      2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

```

```

Qy      22 ERAIKTLNDYELARNECELAGTLFODPSPALPSALGKREAGPY-----SSKTRGMR 74
Db      132 EEEIKLMSPDIEITLYHGESEFGT-----GPIIMISYADESEARVIT 172
Qy      75 WKR-----PFEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLTI 116
Db      173 WKIKIDLPYEVVSTEEKMTKRLRVKEDQPDVLTITNGDNFPAALKRCEKLGVSFTL 232
Qy      117 NEE-ILARVPLNQSFQENYAGIFHPQFQWQGEVVEVVDRLPTKDGELLFVHSAEGSE 175
Db      233 GRDSEPKIQRMGDRFAVEYKGVHFDLYV--IRRTIN--LPYTLLEAVY-----EA 281
Qy      176 FMSALLEKAYAKINGCYEALSGATTGEGFEDTGGIAEW-----YEKKPPEMLFKI 227
Db      282 VFGEKPEKXYAE-----EIAIWEETGEGL--RVARYSMEDARYTYELGR-----ER 326
Qy      228 IQKALQKSLGCSI-DITSADSEAITFOQLVKHAYSVTGAEEVENSGLQKLRIRN 286
Db      327 FPMEAQSLRILQGLMDVRS-----STGNLWEMFLRK 360
Qy      287 PMGEVEMTGRMNDNCPSMNTIDPEERERLTR--HEDGEFW--MSPSDFLRHY 335
Db      361 AYER-----NELAN--KDERELARRGGVAGVYKEPERELMNIYVLDERSLY 409
Qy      336 SRLKIC-NLTPDTLSJDTYKKMLTMDGNMRGSGTAGCCRNYPNTFMNPQYLLIEEB 394
Db      410 PSIIITHNVSPTLNR-----GCRSY----- 431
Qy      395 DEDEDESGSG-----TLVGLQKRRRQKMGEDMHTIGFIVEPEELSGQTN 446
Db      432 DVAPEVGHKPCXDPGFIPLSLGNLLEERQKIRKKKATLIDPEKLLDY-----RQRAI 486
Qy      447 HLSKNFF-----LTNRARESDPTFINLREVLNFKLP----- 478
Db      487 KILANSYGYGYARARWYCRCAESVTAMGEYIEMV--RELEKPFKYLAYDTDGL 544
Qy      479 ----PG-----EYI--LVSTFEPNKGDFCIRVSEKKAQYQAVDEIEA 518
Db      545 HATIPGADAEYKKAAMEFLNYPKLPGLILELEYEG-FYVRGFFVTKKKYAVIDEKGI 603
Qy      519 NLEFPDISDDIDGVRRLFAQLAGEDAISAFELQTIILRVLAQRQDIKSDGFSIETCK 578
Db      604 TTRGLEI-----VRDWSIA--KETQARVLEAILRH-----GDVEAVR 641
Qy      579 IWMVMDLSDSGKLGKHEFYIIMTKIQYQKIRREIDVDSGTMSYEMKALBEAGFKM 638
Db      642 IYREVYER-----LSKYEVEPEKLVHQBQITREIK-DYKATGPEVALAKLAAGVAKI 693
Qy      639 -PCQLHOVIV--ARFADQIIDFNF-----VRCLVLETLFKIF--K 677
Db      694 RPTVVISYIVLKGSGRIGD--RAIPFDEPDPTKHKYDADYIENQVLPAVERILRAFGYR 751
Qy      678 QJDEPENTGTIELDLISWL 695
Db      752 KEDLRVQKTRQVGLQAWL 769

```

```

RESULT 7
US-10-280-962A-46
; Sequence 46, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorige, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity

```

FILE REFERENCE: 25436/2342
CURRENT APPLICATION NUMBER: US/10/280,962A
CURRENT FILING DATE: 2002-10-25
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn version 3.1
SEQ ID NO 46
LENGTH: 776
TYPE: PR
ORGANISM: Thermococcus sp.
US-10-280-962A-46

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

22 ERAIKYLNQDYELALNECLEAGTIFODPSFPAISALGKRELGY-----SKTRGMR 74
132 EEBIKLMSFDIELTYHGEFET-----GPIIMISYADESEARVIT 172
75 WKR-----PTEICADPOEIIIGATRT-----DICOAGLGDCLMLAI-----ASLTL 116
173 WKKIDLPYEVVSTEEKEMIKRPLRVKKEKDPVLITNGDNFPAVLKKGCEKLVSPFL 232
117 NEB-ILARVPLNOSFOENYAGIFHPQFQWQGEWVVDRLPTKDGELLFVHSAEGSE 175
233 GRDSEPKIQRMGDRFAVEVGRVHFDLPV---IRRTIN--LPTVTLVAVY-----EA 281
176 FMSALLEKAYAKINGCEALSGATTEGDFPTGSIAMW-----YELKPPNLFKI 227
282 VFQKPKREKTYAE-----EIAWETGEGLE---RVARYSMEDARVTEIGR-----EF 326
228 IQAKLQKSLGCSI-DITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKLRIRN 286
327 FPMBAQLSRLLIGGLMDVRS-----STGNLVEMWFLRK 360
287 PMGEVWTGRMNDPCSWNTIDPEERELTR-----HEDGEFW--MSFSDFLRHY 335
361 AYER-----NELAPN---KPERELARRGGVAGYVKEPERGLMDINIVYLDPRSLY 409
336 SRLKIC-NLTPDTLSDTYKKMKLTMDGNMRRGSTAGGCRNYPNTFMNPOVYLILBEE 394
410 PSIIITHNVSPDILNRE-----GCRSY-----431
395 DEDEDESGC-----TFVLGLQKRRRQKMGEDMHTTGFQIYEVEBELSGQNTI 446
432 DVAPVGHKFCQDPGFIPLSLGNLBEROKIKRKKATIDPLEKNLDY-----RORAI 486
447 HLSKNF-----LTNRARERSDTFINLREVLNRFKL P-----478
487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELEKFGFKVLYADTDGL 544
479 -----PG-----EYI--LVSTFEPNKGDDFCIRVSEKKADYQAVNDIEIA 518
545 HATIPGADATYKKKAMEFLANTYINPKLPGLLELEYG-FYVRGFPTVKKKYAVIDEGKI 603
519 NLEEFISDEDDIDGVRRLPAOLAGDAEISAFELQTLIRRVLAARODIKSDGSFJETCK 578
604 TTGLEI-----VARDWSEIA--KETQARVLEAILRH-----GDVEAYR 641
579 IAMDMLDSGSGKLGKEFYILMTKIQKYREIDVDRSGTNSYEMRKALEEAGFM 638
642 IYREVTEK-----LSKYEVPEPKLVIHQITRELK-DYKATGPHVALAKRLAARGVKI 693
639 -PCQLHQVIVY---ARFADDQLIIDPDRN-----VRCVLVLETLFKIF--K 677
694 RPTGVISYIVLKSGSGRID--RAIPDEPDPPTKHAKYDADYIENQVLPAYERILARAGYR 751
752 KEDLRKYOKTRQVGLGAWL 769

RESULT 8

US-10-298-680A-45
Sequence 45; Application US/10298680A
GENERAL INFORMATION:
APPLICANT: Sorce, Joseph A
TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
FILE REFERENCE: 25436/2345
CURRENT APPLICATION NUMBER: US/10/298,680A
CURRENT FILING DATE: 2002-11-18
PRIORITY APPLICATION NUMBER: US 10/280,962
PRIORITY FILING DATE: 2002-10-25
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 45
LENGTH: 776
TYPE: PR
ORGANISM: Thermococcus sp.
US-10-298-680A-45

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

22 ERAIKYLNQDYELALNECLEAGTIFODPSFPAISALGKRELGY-----SKTRGMR 74
132 EEBIKLMSFDIELTYHGEFET-----GPIIMISYADESEARVIT 172
75 WKR-----PTEICADPOEIIIGATRT-----DICOAGLGDCLMLAI-----ASLTL 116
173 WKKIDLPYEVVSTEEKEMIKRPLRVKKEKDPVLITNGDNFPAVLKKGCEKLVSPFL 232
117 NEB-ILARVPLNOSFOENYAGIFHPQFQWQGEWVVDRLPTKDGELLFVHSAEGSE 175
233 GRDSEPKIQRMGDRFAVEVGRVHFDLPV---IRRTIN--LPTVTLVAVY-----EA 281
176 FMSALLEKAYAKINGCEALSGATTEGDFPTGSIAMW-----YELKPPNLFKI 227
282 VFQKPKREKTYAE-----EIAWETGEGLE---RVARYSMEDARVTEIGR-----EF 326
228 IQAKLQKSLGCSI-DITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKLRIRN 286
327 FPMBAQLSRLLIGGLMDVRS-----STGNLVEMWFLRK 360
287 PMGEVWTGRMNDPCSWNTIDPEERELTR-----HEDGEFW--MSFSDFLRHY 335
361 AYER-----NELAPN---KPERELARRGGVAGYVKEPERGLMDINIVYLDPRSLY 409
336 SRLKIC-NLTPDTLSDTYKKMKLTMDGNMRRGSTAGGCRNYPNTFMNPOVYLILBEE 394
410 PSIIITHNVSPDILNRE-----GCRSY-----431
395 DEDEDESGC-----TFVLGLQKRRRQKMGEDMHTTGFQIYEVEBELSGQNTI 446
432 DVAPVGHKFCQDPGFIPLSLGNLBEROKIKRKKATIDPLEKNLDY-----RORAI 486
447 HLSKNF-----LTNRARERSDTFINLREVLNRFKL P-----478
487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELEKFGFKVLYADTDGL 544
479 -----PG-----EYI--LVSTFEPNKGDDFCIRVSEKKADYQAVNDIEIA 518
545 HATIPGADATYKKKAMEFLANTYINPKLPGLLELEYG-FYVRGFPTVKKKYAVIDEGKI 603
519 NLEEFISDEDDIDGVRRLPAOLAGDAEISAFELQTLIRRVLAARODIKSDGSFJETCK 578
604 TTGLEI-----VARDWSEIA--KETQARVLEAILRH-----GDVEAYR 641
579 IAMDMLDSGSGKLGKEFYILMTKIQKYREIDVDRSGTNSYEMRKALEEAGFM 638
642 IYREVTEK-----LSKYEVPEPKLVIHQITRELK-DYKATGPHVALAKRLAARGVKI 693
639 -PCQLHQVIVY---ARFADDQLIIDPDRN-----VRCVLVLETLFKIF--K 677
694 RPTGVISYIVLKSGSGRID--RAIPDEPDPPTKHAKYDADYIENQVLPAYERILARAGYR 751

QY 678 QLPDENTGTIELDLISWL 695
Db 752 KEDLRKYQKTRQVGLGAWL 769

RESULT 9
US-10-298-680A-46
; Sequence 46, Application US/10298680A
; GENERAL INFORMATION:
; APPLICANT: Sorige, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2345
; CURRENT APPLICATION NUMBER: US/10/298,680A
; PRIOR FILING DATE: 2002-11-18
; PRIORITY APPLICATION NUMBER: US 10/280,962
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 776
; TYPE: PR1
; ORGANISM: Thermococcus sp.
US-10-298-680A-46

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

QY 22 ERAIKYLNODYEALRECEAGTLFODPSFPAIPALGFKELGPT-----SSKTRGMR 74
Db 132 EEBELKMSFDIELTYHGESEFGT-----GPIIMISYADSEARAVIT 172
QY 75 WKR-----PTEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLT 116
Db 173 WKKIDLPYEVVSTEEKMIRFLRVYKEKDPDLITNGDNFPAIKKRCCKLVGSFTL 232
QY 117 NEB-IILRVVPLNOSFOENYAGIFHPQFOYGEWEVVVDRLPTKDGELLFVHSAEGSE 175
Db 233 GRDSEPKIQMGDRFAVEYKGRVHFDLPV---IRRTIN--LPTTYLLEVY-----EA 281
QY 176 FMSALLEKAYAKINGCEALSGATTEGFEDFTGGIAEW-----YEIKRPPNLFKI 227
Db 282 VFGRPKREKYVAE-----EIAIATWETGEGLR---RVARYSMEDARVTEYELGR-----EF 326
QY 228 IQKALQKSLGCSI-DITSAADSEAITFOKLVKGHAYVTGAEEVESNGSLQKLIRIN 286
Db 327 FPMBAQLSRLIGGLMDVSR-----STGNLVEMWFLLRK 360
QY 287 PMGEVEMTGRMNDNCPSMNTIDPEERERLTR-----HEDGEFW--MSFSDPLRHY 335
Db 361 AYER-----NELAPV---KDERELARRRGVAGYVKEPERGLMDNIIVIDLPSLSY 409
QY 336 SRLIEIC-NLTPDTLTSPTYKKMKLTKMDGNMRGSGTAGCRNYPNTFMNPQYLIKLEEE 394
Db 410 PSIIITHNVSPDILNRE-----GCRSY----- 431
QY 395 DEDEBDESGC-----TFLVGLIOKRRRQKMGEDMHTTGFGIYVPELSGOTNI 446
Db 432 DVAPVGHKFCDFPGFIPSLGNLLEERQKIRKKKATLDPLEKNLLDY-----RORAI 486
QY 447 HLSKNFP-----LTNRASRSDTFINLREVLRFKLP----- 478
Db 487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELBEKFGPKVUYADTDGL 544
QY 479 ----PG-----EYI--LVSTPEPNKGDPCIRVFSSEKKADYQAVDEIEA 518
Db 545 HATIPGADAEIVKKAMEFLANYPKLPGLLELEYEG--FYVRGFFVTKKKAVAVIDECKI 603
QY 519 NLEBFDISEDDIDGVARLFAOLAGDEAIESAFELQITLRLVLAQRDIDISDGSISTETCK 578
Db 604 TTGGLLEI-----VARDWSEIA--KETQARVLEAILRH-----GDVEEAVR 641

QY 579 IMVMDLSDSGSKGLKEFYIITMTKIQKTYREIDVRSCTMNSYEMRKALIEAGFKM 638
Db 642 IYREVEK-----LSKYEVPPEKLIHQBOTREIK-DYKATGPHVALAKLAARGVKI 693
QY 639 -PCQLHGVY-----ARFADQILIIDPDNF-----VRCUNLETLFKIF--K 677
Db 694 RFGTVISYIVLKSSGSGIGD--RAIPDEDPPTKHYDADYIENQVLAVERILRAFGRY 751
QY 678 QLPDENTGTIELDLISWL 695
Db 752 KEDLRKYQKTRQVGLGAWL 769

RESULT 10
US-10-324-846-55
; Sequence 55, Application US/10324846
; GENERAL INFORMATION:
; APPLICANT: Bohns, Michael
; TITLE OF INVENTION: DNA Polymerase Blends and Uses Thereof
; FILE REFERENCE: 25436/2362
; CURRENT APPLICATION NUMBER: US/10/324,846
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 776
; TYPE: PR1
; ORGANISM: Pyrococcus GB-D
US-10-324-846-55

Query Match 2.8%; Score 105; DB 6; Length 776;
Best Local Similarity 18.9%; Pred. No. 0.16;
Matches 151; Conservative 105; Mismatches 258; Indels 284; Gaps 40;

QY 22 ERAIKYLNODYEALRECEAGTLFODPSFPAIPALGFKELGPT-----SSKTRGMR 74
Db 132 EEBELKMSFDIELTYHGESEFGT-----GPIIMISYADSEARAVIT 172
QY 75 WKR-----PTEICADPOFIIGATRT-----DICQALGDCWLLAI-----ASLT 116
Db 173 WKKIDLPYEVVSTEEKMIRFLRVYKEKDPDLITNGDNFPAIKKRCCKLVGSFTL 232
QY 117 NEB-IILRVVPLNOSFOENYAGIFHPQFOYGEWEVVVDRLPTKDGELLFVHSAEGSE 175
Db 233 GRDSEPKIQMGDRFAVEYKGRVHFDLPV---IRRTIN--LPTTYLLEVY-----EA 281
QY 176 FMSALLEKAYAKINGCEALSGATTEGFEDFTGGIAEW-----YEIKRPPNLFKI 227
Db 282 VFGRPKREKYVAE-----EIAIATWETGEGLR---RVARYSMEDARVTEYELGR-----EF 326
QY 228 IQKALQKSLGCSI-DITSAADSEAITFOKLVKGHAYVTGAEEVESNGSLQKLIRIN 286
Db 327 FPMBAQLSRLIGGLMDVSR-----STGNLVEMWFLLRK 360
QY 287 PMGEVEMTGRMNDNCPSMNTIDPEERERLTR-----HEDGEFW--MSFSDPLRHY 335
Db 361 AYER-----NELAPV---KDERELARRRGVAGYVKEPERGLMDNIIVIDLPSLSY 409
QY 336 SRLIEIC-NLTPDTLTSPTYKKMKLTKMDGNMRGSGTAGCRNYPNTFMNPQYLIKLEEE 394
Db 410 PSIIITHNVSPDILNRE-----GCRSY----- 431
QY 395 DEDEBDESGC-----TFLVGLIOKRRRQKMGEDMHTTGFGIYVPELSGOTNI 446
Db 432 DVAPVGHKFCDFPGFIPSLGNLLEERQKIRKKKATLDPLEKNLLDY-----RORAI 486
QY 447 HLSKNFP-----LTNRASRSDTFINLREVLRFKLP----- 478
Db 487 KILANSYGYGYARARAWYCRECAESVTAMGREYIEMVI--RELBEKFGPKVUYADTDGL 544
QY 479 ----PG-----EYI--LVSTPEPNKGDPCIRVFSSEKKADYQAVDEIEA 518
Db 545 HATIPGADAEIVKKAMEFLANYPKLPGLLELEYEG--FYVRGFFVTKKKAVAVIDECKI 603

```

QY      519 NLEBPDISDDIDDGVRRLPFAQLAGDAEISAFELQTIIRRYLAKRQDIXKDSGFSIETCK 578
Db      604 TTRGCEI-----VRDMSEIA--KETOARVLEAILRH-----GDVEAYR 641
QY      579 IMVMDLSDSGSKLGKKEFYILMTKIQXYQKIYREIDVDRSGTMSYEMRKALEAGFYK 638
Db      642 IYREVTYK-----LSKYEVPPEKLVIEHQITRELK-DYKATGPHVALAKRLLAAGVKI 693
QY      639 -PCQHQYIV---ARFADDQLIDPDN-----VRLVLETLTKIE-K 677
Db      694 RRGVTSIYVLKSGSGIGD-RAIPDEFDPYTKKDYADYIENOVLPAYERILRAFGYR 751
QY      678 QUDPENTGTIEDLISWL 695
Db      752 KEDLRYQKTRQVGLGAWL 769

RESULT 11
US-10-324-846-56
: Sequence 56, Application US/10324846
: GENERAL INFORMATION:
:   APPLICANT: Borns, Michael
:   TITLE OF INVENTION: DNA Polymerase Blends and Uses Thereof
:   FILE REFERENCE: 25436/2362
:   CURRENT APPLICATION NUMBER: US/10/324,846
:   CURRENT FILING DATE: 2002-12-20
:   NUMBER OF SEQ ID NOS: 113
:   SOFTWARE: PatentIn version 3.1
: SEQ ID NO 56
:   LENGTH: 776
:   TYPE: PRT
:   ORGANISM: Pyrococcus GB-D
: US-10-324-846-56

```

Query Match	2.8%;	Score 105;	DB 6;	Length 776;
Best Local Similarity	18.9%;	Pred. No. 0.16;		
Matches 151;	Conservative 105;	Mismatches 258;	Indels 284;	Gaps 40

```

0Y 22 ERAIYVNDYVYALNRECEAGLTQDPSPFAIPALGFKEGPPY-----SSKTRGMR 74
Db 132 EBEILKMSDIDETLHGESEFGT-----GPIIMISVADSEARVIT 172
QY 75 WKR---PFEICADPOFIIGATRT---DICOGLADPCWLLAAI-----ASLTL 116
Db 173 WKKIDLPYVEVSTEKEMIKRFLRVYKEXODPVLTITNGDNFDPFALYKRCCKELGVSTL 232
QY 117 NEE-LLARVYPLNQSFOENYAGTFHQFOQYGEWVEVVDRLPTKDGELLFPHSAGSE 175
Db 233 GRDGSEPKIQRMGDRPAVEKRVHVDLVPV-----IRRTIN--LPTVTLLEVY-----EA 281
QY 176 FWSALLERKAYAKINGCYEALSGATTEGEFDTGGIAEW-----YELKKEPPNLFKI 227
Db 282 VFGKREKEKYAE-----EIAIATMEGBE---RARVSMEDARVITYELGR-----EF 326
QY 228 IOKALQKSGLLGCSI--DITSADSEAITFOKLVGHAYSVTGAEEVESNGSLQKLRIRN 286
Db 327 FPMEAQLSRLLIQGLMDVRSR-----STGNLYEWFLLRK 360
QY 287 PMGEVEMTGRNMNDNCMSMTIDPEERELTR-----HEDGEFW--MSSSDPLRHY 335
Db 361 AER-----NELAPN---KDERELARRRGVAGGYKEBERGLMDNIIVLDPSLY 409
QY 336 SRLIEC-NUPTPTLTSDTYKKMKLTMQDGNMRGSTAGGRANYPNTFMNPOYLKLEE 394
Db 410 PSIIITTHVSPTLNRE-----GCRSY----- 431
QY 395 DEDEEDGESG-----TFLVGLIQKHRRORKMGEDMTATIGIYVEBEELSQTNI 446
Db 432 DVAPFVGHKFCDFGFIPLSLIGNLIEEROKIKRKKAKATLDPLEKUNLDY-----QORAI 486
QY 447 HLSKRF-----LTRARARSSTFINLRVLRFKLP----- 478

```

```

Dh 487 KILANYYGYGYABARWYRECACESVYAMOREVIEMWY--RELBEKFGFEKVIYADBDGL 544
Qy 479 -----PG-----EYI--LVSTSEPKKDDGFCIRVSEKKADYQAVNDEIEA 51.8
      |||
Dh 545 HATIPADAEYKAKKAMEFLNINPKLGGILLEYEYG--FYVKGFFYTKKXAYVIDEGKI 603
      |||
Qy 519 NNEEFDISEDIDDDGVRRFLPAQAGEDAEISAFELQTLIRRVLAKRODIXSDGFSIETCK 578
      |||
Dh 604 TTRGTEI-----VRDRWSEIHA--KETQARVLEALIRH-----GDVEEAVR 641
      |||
Qy 579 IAMDMLDSDSGSKLGLKEFYILMTQYKQYVREIYDVDRSGTNNYSYEMERKALEEAEFKM 63.8
      |||
Dh 642 IYAEVIEYK-----LSKTEVPEPEGVHGOITBELK--DYKATPPhYAIKRLAARGVKI 653
      |||
Qy 639 -PCQLHQVIV---ARFADQLIIDPNF-----VRCLVRLETLFEKIF--K 677
      |||
Dh 694 RFGTVISYIVLKGSGRIGD--RAIPDFEDEFPTKKHKYADYIENQVPAVERILIRAGYR 751
      |||
Qy 678 QLDPEYGTIELDLISWL 695
      |||
Dh 752 KEDLRYYOKTRVGLGAWL 769
      |||

```

```

RESULT 12
US-10-280-962A-41
; Sequence 41, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorce, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2342
; CURRENT APPLICATION NUMBER: US/10/280,962A
; CURRENT FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 774
; TYPE: PR1
; ORGANISM: Pyrococcus sp.
US-10-280-962A-41

```

Query Match	2.8#:	Score	10.4.5:	DB 6:	Length	774:
Best Local Similarity	19.7#:	Pred.	0.17#:			
Matches 131:	Conservative 104:	Mismatches 238:	Indels 191:	Gaps	33	
Qy	149	WVEVVDNR-----	LPTKDELLFVHSAGSEFPWSALLEKXAVAKINGCYEALSGA	199		
Db	180	YVDVASTEREMIKRFLRVYKEKDDVLTYYGNDF--	ALYKRCETL-GINFALGRDG	236		
Qy	200	TTESEFDPTGSIARWYELK-KEPNLFKTIOKALQKSLGSCSIDTSAADSEATFPKL	258			
Db	237	SEPKIQRMGRFA--VEYGRIHDFIVIRITNLP-	YTLAEVYEAVFQPK	287		
Qy	259	VKGHAYSVTGAEEVSN-----	GSLOLIRIRNFWGVE	292		
Db	288	EKYAAEETIPAMETGENLERVARYSMEDAKTYELGKEFLPMEAOISRLI	CGSL	341		
Qy	293	W-----TG--RW-----	NDNCQSNWTIDPEERERLTRRHEDE-----	F	324	
Db	342	WDVRRSSTGNLVFPLFLKRAYERELAN----	KDEKE-LARRQSYEGGYKEPERGL	396		
Qy	325	W--MSFDPLFRHRYRLIC-NLTPDPLTSDTYKKMKLTKMDGNMRRGSTAGCRNYPNTF		381		
Db	397	WENIVYDFRSLYSLIITHNVSDPLNRECKCEYDVAPOUGH-----	RFCRDFPGFI	449		
Qy	382	WMNPQYL-IKLEEDDEDEDESGCTFLVGLIOKHRRQRQKMGEMHTIGG-		432		
Db	450	---PSLLGDLLEEROKIKKKKATIDPIERLLDYRQCAITILANSYYGYGYARAWYC	506			
Qy	433	-----IYVPEELSGOTNIHLSKNFLLTNRAERSDTEINLEEVNLR	474			
Db	507	KECAESVTAMGREYITMTIKIEIKKYGKVIYSDTDGFATIPGADAFTVKKKAEPFN-	565			
Qy	475	FKLPGREYI--LVSTEEPNKDGDFCI RVSEBKADYQAVUDEIFANLEEFPISSDDIDD	532			

Db 566 -----YINAKLPGALELEYEG-FYKRGFFVTKKKAVIIDESEKITTRGLEI----- 610
Qy 533 GVRRLPAQLAGDAEISAFELQTLIRVLAKRODICKSDGFSIETCKIMVMDLSDSGKL 592
Db 611 -VRDMSSEA---KETQARVLBALK-----DGVKAVRIVKEVTEK----- 649
Qy 593 GLKEFYILMTKIOKQKIYREIDVDSGTMSYEMKALBEAGFKM-PCQLHGVIV---- 647
Db 650 -LSKYVPEPKLVIHQITRDLC-DYKATGPHVAVKRLAARGVKIRPGTVISYIVLKS 707
Qy 648 ARPADQLIIDPFNF-----VRCLVLETLFKIF--KQDPENGTIYLDL 691
Db 708 GRIGD--RAIPFBEFPTKHKYDAEYIENQVPAVERIIRAFGRKEDIRYOKTRQVGL 765
Qy 692 ISWL 695
Db 766 SAWL 769

RESULT 13

US-10-280-962A-42
; Sequence 42, Application US/10280962A
; GENERAL INFORMATION:
; APPLICANT: Sorage, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2342
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US/10/280, 962A
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 774
; TYPE: PRF
; ORGANISM: Pyrococcus sp.
US-10-280-962A-42

Query Match 2.8%; Score 104.5; DB 6; Length 774;
Best Local Similarity 19.7%; Pred. No. 0.17;
Matches 131; Conservative 104; Mismatches 238; Indels 191; Gaps 33;

Qy 149 WVEYVVDNR-----LPTKDGELLFVHSAEGSEFWALLKAYAKINGCYEALSOGA 199
Db 180 YDVVSTEREMIKRFLRVKEXKDPDVLITYNGDNFDF--AYLKRCCEKL-GINFALGRDG 236
Qy 200 TTGEFEDFTGGLAEWEYLK-KPPNLFKIIQKALQKSGSLGCSIDITSADSEAITFOKL 258
Db 237 SEPKIQMGDRFA--VEVKGRIHFDLYPVIRRTINLPT-----YTLAEVYEAVFGQPK 287
Qy 259 VKGHAYSVTGAEEVESN-----GSLQKLRIRNPMGEVE 292
Db 288 EKYYAEIITPAWETGENLERVARYSMEDAKVYELGKEFLPMHQSLRLI-----GQSL 341
Qy 293 W-----TG--RW-----NDNCPSMNTIDPEERELTRRHEDG-----F 324
Db 342 MDVSRSGTGNLVEMFLRKAYERNELAPN---KPEKE-LARRQSYEGGYVEBERGL 396
Qy 325 W--MSFSDFLRHSRLKIC-NLPDPTLSDTYKKMKLTKMDGNMRSGSTAGGRNPNYF 381
Db 397 WENIVYLDPRSLPSTIITHNVSPTLNREGCKEYDAPQVGH-----RFCKDPGFT 449
Qy 382 WNNPOYL-ITLBEDEDEBDESGCTFLVGLIQHRRRORRKMGEDEMTTIGFG----- 432
Db 450 ---PSLLGDLLEBRQKIKKKMKATIDPIERKLDYGRRAIKILANSYGYGYARARWYC 506
Qy 433 -----IYVPEELSGQTNILHSKNFPLTNRARSSTFINLREVLNR 474
Db 507 KECAESVTAMGREYITMTIKIEEKYGFYISDTPGFAITPGADAETVKKKAMEFLN- 565
Qy 475 FKLPGBEYI--LVPSFTEPNKGDGFCIRVSEKKADYQAVDEIEANLEBFDISEDDIDD 532
Db 566 -----YINAKLPGALELEYEG-FYKRGFFVTKKKAVIIDESEKITTRGLEI----- 610

Qy 533 GVRRLPAQLAGDAEISAFELQTLIRVLAKRODICKSDGFSIETCKIMVMDLSDSGKL 592
Db 611 -VRDMSSEA---KETQARVLBALK-----DGVKAVRIVKEVTEK----- 649
Qy 593 GLKEFYILMTKIOKQKIYREIDVDSGTMSYEMKALBEAGFKM-PCQLHGVIV---- 647
Db 650 -LSKYVPEPKLVIHQITRDLC-DYKATGPHVAVKRLAARGVKIRPGTVISYIVLKS 707
Qy 648 ARPADQLIIDPFNF-----VRCLVLETLFKIF--KQDPENGTIYLDL 691
Db 708 GRIGD--RAIPFBEFPTKHKYDAEYIENQVPAVERIIRAFGRKEDIRYOKTRQVGL 765
Qy 692 ISWL 695
Db 766 SAWL 769

RESULT 14

US-10-298-680A-41
; Sequence 41, Application US/10298680A
; GENERAL INFORMATION:
; APPLICANT: Sorage, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2345
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US 10/280, 962
; PRIOR FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 774
; TYPE: PRF
; ORGANISM: Pyrococcus sp.
US-10-298-680A-41

Query Match 2.8%; Score 104.5; DB 6; Length 774;
Best Local Similarity 19.7%; Pred. No. 0.17;
Matches 131; Conservative 104; Mismatches 238; Indels 191; Gaps 33;

Qy 149 WVEYVVDNR-----LPTKDGELLFVHSAEGSEFWALLKAYAKINGCYEALSOGA 199
Db 180 YDVVSTEREMIKRFLRVKEXKDPDVLITYNGDNFDF--AYLKRCCEKL-GINFALGRDG 236
Qy 200 TTGEFEDFTGGLAEWEYLK-KPPNLFKIIQKALQKSGSLGCSIDITSADSEAITFOKL 258
Db 237 SEPKIQMGDRFA--VEVKGRIHFDLYPVIRRTINLPT-----YTLAEVYEAVFGQPK 287
Qy 259 VKGHAYSVTGAEEVESN-----GSLQKLRIRNPMGEVE 292
Db 288 EKYYAEIITPAWETGENLERVARYSMEDAKVYELGKEFLPMHQSLRLI-----GQSL 341
Qy 293 W-----TG--RW-----NDNCPSMNTIDPEERELTRRHEDG-----F 324
Db 342 MDVSRSGTGNLVEMFLRKAYERNELAPN---KPEKE-LARRQSYEGGYVEBERGL 396
Qy 325 W--MSFSDFLRHSRLKIC-NLPDPTLSDTYKKMKLTKMDGNMRSGSTAGGRNPNYF 381
Db 397 WENIVYLDPRSLPSTIITHNVSPTLNREGCKEYDAPQVGH-----RFCKDPGFT 449
Qy 382 WNNPOYL-ITLBEDEDEBDESGCTFLVGLIQHRRRORRKMGEDEMTTIGFG----- 432
Db 450 ---PSLLGDLLEBRQKIKKKMKATIDPIERKLDYGRRAIKILANSYGYGYARARWYC 506
Qy 433 -----IYVPEELSGQTNILHSKNFPLTNRARSSTFINLREVLNR 474
Db 507 KECAESVTAMGREYITMTIKIEEKYGFYISDTPGFAITPGADAETVKKKAMEFLN- 565
Qy 475 FKLPGBEYI--LVPSFTEPNKGDGFCIRVSEKKADYQAVDEIEANLEBFDISEDDIDD 532
Db 566 -----YINAKLPGALELEYEG-FYKRGFFVTKKKAVIIDESEKITTRGLEI----- 610
Qy 533 GVRRLPAQLAGDAEISAFELQTLIRVLAKRODICKSDGFSIETCKIMVMDLSDSGKL 592

Db 611 -VRDWSBIA---KETOQARVLEALLK-----DGDVEKAVRIKVEYTER----- 649
Qy 593 GLKEFYILMTKIQKQKTYREIDVDRSGTNNSEYMRKALEBAGFKM-PCQLHQVIV----- 647
Db 650 -LSKTYEVPPEKLVHIEQITRDLK-DYKATGPHVAVAKRLAARGVKIRPGVIVSYIVLKG 707
Qy 648 ARFADQOLIIDFNF-----VRCLVLETLEFKIF--KQDPENTGTIELDL 691
Db 708 GRID--RAIPFDEPTKHKYDAEYIENQVLAVERILRAFGRKEDLRYQKTRQVGL 765
Qy 692 ISWL 695
Db 766 SAWL 769

RESULT 15
US-10-298-680A-42
; Sequence 42, Application US/10298680A
; GENERAL INFORMATION:
; APPLICANT: Sarge, Joseph A
; TITLE OF INVENTION: DNA Polymerases with Reduced Base Analog Detection Activity
; FILE REFERENCE: 25436/2345
; CURRENT APPLICATION NUMBER: US/10/298, 680A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US 10/280, 962
; PRIOR FILING DATE: 2002-10-25
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Pyrococcus sp.
US-10-298-680A-42

Query Match 2.8%; Score 104.5; DB 6; Length 774;
Best Local Similarity 19.7%; Pred. No. 0.17;
Matches 131; Conservative 104; Mismatches 238; Indels 191; Gaps 33;

Qy 149 WVEVVVDDR-----LPTKDGELLFVHSAEGSEFWSALLEKAYAKINGCYEALSGGA 199
Db 180 YDVVSTREMIKRELVVKEKEDPVLITNGDNDF--AYIKKRCERL-GINFALGRDG 236
Qy 200 TTEGFEDFTGGAIEWYELK-KPPNLFKTIQALOKGSLGCSIDITSADSEAITFOKL 258
Db 237 SEPKIQRMGDRFA--VEYKGRHFDLPYVIRITNLP-----YTLEAVYEAVFGQPK 287
Qy 259 VKGHAYSVTGAEEVESN-----GSLQKLRIRNFWGEVE 292
Db 288 EKVAABEITPAMETGENLERVARYSMEDAKTYELGKEFLPMEQLSRLI-----GQSL 341
Qy 293 W-----TG---RW-----NDNCPSWNTIDPEERELTRHEDG-----F 324
Db 342 WDVSRSSGTNLVEMFLKRAYERNELAPN---KPEKE-LARRQSYEGGVXKEPERGL 396
Qy 325 W--MSFSDPLRHSYLEIC-NLTPTLTSDTYKWKLTMDGNMRSGTAGCENYPTF 381
Db 397 WENIVYLDPRSLYPSIIITTHVSPDLNREGCKEYDVAPOYGH-----RCKDFPGFI 449
Qy 382 WNPQYL-IKLEEDDEDEDESGCTFLVGLIQHRRRQKXGEMDTIGF----- 432
Db 450 ---PSLGDLLERQKIKKKMKATIDPIERKLDYQRAIKILANSYGYGYAVARWYC 506
Qy 433 -----IYVPELSGOTNIHLSKNFLLTNRAERSDTPINLREVNR 474
Db 507 KCAESVTAMGREYITMTIKETIEKYGRVIYSDTGFFATIPGADAEIVKKAMEFLN- 565
Qy 475 FKLPGGEYI--LVPEFPNKDGDRCIRVSEKADYOAVDDEIRANLEEDFIDSEDIDD 532
Db 566 -----YINAKLPALALEYEG--FYKGFYTKKKYAVIDEGKITRGLEI----- 610
Qy 533 GVRRLFAQLAGEDAEISAFELQTLIRVLARQDIKSDGFSIETCKIMVMDLSDSGSL 592

Db 611 -VRDWSBIA---KETOQARVLEALLK-----DGDVEKAVRIKVEYTER----- 649
Qy 593 GLKEFYILMTKIQKQKTYREIDVDRSGTNNSEYMRKALEBAGFKM-PCQLHQVIV----- 647
Db 650 -LSKTYEVPPEKLVHIEQITRDLK-DYKATGPHVAVAKRLAARGVKIRPGVIVSYIVLKG 707
Qy 648 ARFADQOLIIDFNF-----VRCLVLETLEFKIF--KQDPENTGTIELDL 691
Db 708 GRID--RAIPFDEPTKHKYDAEYIENQVLAVERILRAFGRKEDLRYQKTRQVGL 765
Qy 692 ISWL 695
Db 766 SAWL 769

Search completed: July 24, 2003, 13:13:09
Job time : 11.7077 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:51:55 ; Search time 422.806 Seconds

(without alignments)
1440.970 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713
Sequence: 1 MAGIAAKLAKDRRAEGLS.....PENTGTIEDLISWLCFSVL 700

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5580241 segs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1: /cgn2_6/prodata/1/paa/PCTUS_COMB.pep:*
2: /cgn2_6/prodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/prodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/prodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/prodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/prodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/prodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/prodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/prodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/prodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/prodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/prodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/prodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/prodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/prodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/prodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/prodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/prodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/prodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/prodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/prodata/1/paa/US097A_COMB.pep:*
22: /cgn2_6/prodata/1/paa/US097B_COMB.pep:*
23: /cgn2_6/prodata/1/paa/US098_COMB.pep:*
24: /cgn2_6/prodata/1/paa/US099A_COMB.pep:*
25: /cgn2_6/prodata/1/paa/US099B_COMB.pep:*
26: /cgn2_6/prodata/1/paa/US100_COMB.pep:*
27: /cgn2_6/prodata/1/paa/US101_COMB.pep:*
28: /cgn2_6/prodata/1/paa/US102_COMB.pep:*
29: /cgn2_6/prodata/1/paa/US103_COMB.pep:*
30: /cgn2_6/prodata/1/paa/US104_COMB.pep:*
31: /cgn2_6/prodata/1/paa/US106_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3713	100.0	700	1	PCT-US02-10419-12
2	3713	100.0	700	16	US-09-281-253-12

ALIGNMENTS

RESULT 1
PCT-US02-10419-12
Sequence 12, Application PC/TUS0210419
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
FILE REFERENCE: D0124 PCT
CURRENT APPLICATION NUMBER: PCT/US02/10419
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: US 60/281,253
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/288,768
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/296,180
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/300,620
PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-10419-12

3	3713	100.0	700	17	US-09-301-274-9	Sequence 9, Appl1
4	3713	100.0	700	19	US-09-538-092-947	Sequence 947, Appl
5	3713	100.0	700	19	US-09-562-979-9	Sequence 9, Appl
6	3713	100.0	700	22	US-09-768-877-23	Sequence 23, Appl
7	3713	100.0	700	22	US-09-791-537-8310	Sequence 81010, A
8	3713	100.0	700	23	US-09-840-707A-9	Sequence 9, Appl1
9	3713	100.0	700	23	US-09-840-707B-9	Sequence 9, Appl1
10	3713	100.0	700	23	US-09-884-319A-7	Sequence 7, Appl1
11	3713	100.0	700	26	US-10-038-557A-9	Sequence 9, Appl1
12	3713	100.0	700	27	US-10-116-519-12	Sequence 12, Appl
13	3713	100.0	700	28	US-10-219-051B-12485	Sequence 12485, A
14	3713	100.0	700	28	US-10-293-017-78	Sequence 78, Appl
15	3713	100.0	700	28	US-10-293-071-78	Sequence 78, Appl
16	3713	100.0	700	31	US-10-293-181-12	Sequence 12, Appl
17	3713	100.0	700	31	US-10-296-670-178	Sequence 12, Appl
18	3713	100.0	700	31	US-10-300-620-12	Sequence 4, Appl
19	3713	100.0	700	31	US-10-350-064-4	Sequence 390, App
20	3713	100.0	700	31	US-10-389-987-390	Sequence 390, App
21	3713	100.0	700	31	US-10-412-418-390	Sequence 2, Appl1
22	3705	99.8	700	21	US-09-736-860-2	Sequence 7, Appl1
23	3705	99.8	700	21	US-09-736-860-7	Sequence 9, Appl1
24	3705	99.8	700	21	US-09-736-860-9	Sequence 63859, A
25	3705	99.8	700	22	US-09-791-537-63859	Sequence 63859, A
26	3698	99.6	699	22	US-09-791-537-65839	Sequence 65839, A
27	3698	99.6	699	31	US-10-350-064-96	Sequence 96, Appl
28	3666.5	98.7	701	31	US-10-443-556-2373	Sequence 2373, Ap
29	3666.5	98.7	701	31	US-10-452-680-23239	Sequence 23239, Ap
30	3666.5	98.7	701	31	US-10-453-050-14515	Sequence 14515, A
31	3666.5	98.7	701	31	US-10-453-135-14515	Sequence 14515, A
32	3666.5	98.7	701	31	US-10-455-444-7816	Sequence 7816, Ap
33	3666.5	98.7	701	31	US-10-465-241-7816	Sequence 7815, Ap
34	3666.5	98.7	701	31	US-10-466-412-14515	Sequence 14515, A
35	3657	98.5	754	21	US-09-724-678-81227	Sequence 81227, A
36	3657	98.5	754	21	US-09-724-678A-81227	Sequence 81227, A
37	3522	94.9	700	22	US-09-791-537-122174	Sequence 122174, A
38	3518	94.7	700	21	US-09-736-860-4	Sequence 4, Appl1
39	3518	94.7	700	21	US-09-736-860-8	Sequence 8, Appl1
40	3518	94.7	700	22	US-09-791-537-91444	Sequence 91444, A
41	3518	94.7	700	22	US-09-791-537-131857	Sequence 131857, A
42	3518	94.7	700	31	US-10-350-064-12	Sequence 12, Appl
43	3511	94.6	700	22	US-09-791-537-99714	Sequence 99714, A
44	3417	92.0	691	21	US-09-724-678-81234	Sequence 81234, A
45	3417	92.0	691	21	US-09-724-678A-81234	Sequence 81234, A

Query Match 100.0%; Score 3713; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGIAAKIAKREAAEGISHERAIKYINODYALRNECEAGTLPDPSFPALPSALGF 60
DB 1 MAGIAAKIAKREAAEGISHERAIKYINODYALRNECEAGTLPDPSFPALPSALGF 60
QY 61 KELGPSKTRGMRKRPTEICADPOFTIGATRTDICOALGDCWLLAALASTLINEI 120
DB 61 KELGPSKTRGMRKRPTEICADPOFTIGATRTDICOALGDCWLLAALASTLINEI 120
QY 121 LARVPLNOSFOENVAGIFHFQFQYGEWEVVVDRLPTKGELLFVHABSEFWMSAL 180
DB 121 LARVPLNOSFOENVAGIFHFQFQYGEWEVVVDRLPTKGELLFVHABSEFWMSAL 180
QY 181 LEKAVAKINGCYEALSGATTEGFEPTGGIAEWELCKPPNLFKIIQKALQKSLIGC 240
DB 181 LEKAVAKINGCYEALSGATTEGFEPTGGIAEWELCKPPNLFKIIQKALQKSLIGC 240
QY 241 SIDITSAADSEALTTFQKLVKGHAYSVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
DB 241 SIDITSAADSEALTTFQKLVKGHAYSVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
QY 301 CPSMNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
QY 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBESGCTFLVGLIQKRRROR 420
DB 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBESGCTFLVGLIQKRRROR 420
QY 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPBG 480
DB 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPBG 480
QY 481 EYILVSTFEENKGDGFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
DB 481 EYILVSTFEENKGDGFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKKEFYIL 600
QY 601 WTKIQOKYKRIEIDVRSSTWNSYEMRKALBEAGFMPQQLHOVIVARFADQLIIDPD 660
DB 601 WTKIQOKYKRIEIDVRSSTWNSYEMRKALBEAGFMPQQLHOVIVARFADQLIIDPD 660
QY 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
```

RESULT 2

US-09-281-253-12
; Sequence 12, Application US/09281253
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124
; CURRENT APPLICATION NUMBER: US/09/281,253
; CURRENT FILING DATE: 2001-04-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-281-253-12

Query Match 100.0%; Score 3713; DB 16; Length 700;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MAGIAALADREAAEGISHERAIKYINODYALRNECEAGTLPDPSFPALPSALGF 60
DB 1 MAGIAALADREAAEGISHERAIKYINODYALRNECEAGTLPDPSFPALPSALGF 60
QY 61 KELGPSKTRGMRKRPTEICADPOFTIGATRTDICOALGDCWLLAALASTLINEI 120
DB 61 KELGPSKTRGMRKRPTEICADPOFTIGATRTDICOALGDCWLLAALASTLINEI 120
QY 121 LARVPLNOSFOENVAGIFHFQFQYGEWEVVVDRLPTKGELLFVHABSEFWMSAL 180
DB 121 LARVPLNOSFOENVAGIFHFQFQYGEWEVVVDRLPTKGELLFVHABSEFWMSAL 180
QY 181 LEKAVAKINGCYEALSGATTEGFEPTGGIAEWELCKPPNLFKIIQKALQKSLIGC 240
DB 181 LEKAVAKINGCYEALSGATTEGFEPTGGIAEWELCKPPNLFKIIQKALQKSLIGC 240
QY 241 SIDITSAADSEALTTFQKLVKGHAYSVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
DB 241 SIDITSAADSEALTTFQKLVKGHAYSVTGAEEVNSGSLQKLIIRNPMGEVEMTGRMND 300
QY 301 CPSMNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPTLTSTYKKWKLTK 360
QY 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBESGCTFLVGLIQKRRROR 420
DB 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEEBEDEBESGCTFLVGLIQKRRROR 420
QY 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPBG 480
DB 421 KMGEDMHTIGFGIYEVEPELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPBG 480
QY 481 EYILVSTFEENKGDGFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
DB 481 EYILVSTFEENKGDGFCIRVSEKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAKRQDIKSGFSIETCKIMVMDLSDSGSKGLKKEFYIL 600
QY 601 WTKIQOKYKRIEIDVRSSTWNSYEMRKALBEAGFMPQQLHOVIVARFADQLIIDPD 660
DB 601 WTKIQOKYKRIEIDVRSSTWNSYEMRKALBEAGFMPQQLHOVIVARFADQLIIDPD 660
QY 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
```

RESULT 3

US-09-301-274-9
; Sequence 9, Application US/09301274
; GENERAL INFORMATION:
; APPLICANT: Frederick, Terry M.
; APPLICANT: Ignatkev, George M.
; APPLICANT: ANTIBODY SYSTEMS, INC.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; FILE REFERENCE: 24881-0301
; CURRENT APPLICATION NUMBER: US/09/301,274
; CURRENT FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: IL-1 receptor intracellular ligand protein
; OTHER INFORMATION: comprising amino acid sequence

;; PUBLICATION INFORMATION:
;; PATENT DOCUMENT NUMBER: 5,817,476
;; PATENT FILING DATE: 1995-06-07
;; PUBLICATION DATE: 1998-10-06
US-09-301-274-9

Query Match 100.0%; Score 3713; DB 17; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAGIAATLADREABEGSGHERAIKYLNDYELARNCEAGTLFDDPSFPAISALGF 60
DB 1 MAGIAATLADREABEGSGHERAIKYLNDYELARNCEAGTLFDDPSFPAISALGF 60
QY 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICQALGDCWLLAIASTLTNEEI 120
DB 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICQALGDCWLLAIASTLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFWQGEWVEVVDRLPTKDBELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFWQGEWVEVVDRLPTKDBELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYLKKRPPNLFKIIQKALQKSLGC 240
DB 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYLKKRPPNLFKIIQKALQKSLGC 240
QY 241 SIITSAADSEALTFOKLIVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
DB 241 SIITSAADSEALTFOKLIVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
QY 301 CPSWNTIDPEBERRLTRRHEDGEFWSFSDFLRHYSRLIEICNLTPDTLTSPTYKKMLTK 360
DB 301 CPSWNTIDPEBERRLTRRHEDGEFWSFSDFLRHYSRLIEICNLTPDTLTSPTYKKMLTK 360
QY 361 MDGNWRGSAAGCGRNTPNTFMNPOYLKLEEBDEDEBEGSGCTFLVGLIQHRRROR 420
DB 361 MDGNWRGSAAGCGRNTPNTFMNPOYLKLEEBDEDEBEGSGCTFLVGLIQHRRROR 420
QY 421 KMGEDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNARERSDPTFLNREVLNRFKLPBG 480
DB 421 KMGEDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNARERSDPTFLNREVLNRFKLPBG 480
QY 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVNDEIEANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVNDEIEANLEEFDISDDIDGVRRLFAQ 540
QY 541 LAGSDAISAPELQTIIRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGIKEFYIL 600
DB 541 LAGSDAISAPELQTIIRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGIKEFYIL 600
QY 601 WTIKIQOKIYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
DB 601 WTIKIQOKIYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
QY 661 NFVRCIVRLLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCIVRLLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

```

RESULT 4
US-09-538-092-947
; Sequence 947, Application US/09538092
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; PRIOR APPLICATION NUMBER: 2000-03-29
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/117,965
; PRIOR FILING DATE: 2000-02-01

;; NUMBER OF SEQ ID NOS: 1387
;; SOFTWARE: CuraFastSeqFormatter Version 0.9
;; SEQ ID NO 947
;; LENGTH: 700
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P17655
US-09-538-092-947

Query Match 100.0%; Score 3713; DB 19; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAGIAATLADREABEGSGHERAIKYLNDYELARNCEAGTLFDDPSFPAISALGF 60
DB 1 MAGIAATLADREABEGSGHERAIKYLNDYELARNCEAGTLFDDPSFPAISALGF 60
QY 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICQALGDCWLLAIASTLTNEEI 120
DB 61 KELGPYSKTRGMRKRPTEICADPOFIIGATRTDICQALGDCWLLAIASTLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFWQGEWVEVVDRLPTKDBELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFWQGEWVEVVDRLPTKDBELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYLKKRPPNLFKIIQKALQKSLGC 240
DB 181 LERAYAKINGCYEALSGATTEGFEDEFTGIAEWEYLKKRPPNLFKIIQKALQKSLGC 240
QY 241 SIITSAADSEALTFOKLIVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
DB 241 SIITSAADSEALTFOKLIVGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
QY 301 CPSWNTIDPEBERRLTRRHEDGEFWSFSDFLRHYSRLIEICNLTPDTLTSPTYKKMLTK 360
DB 301 CPSWNTIDPEBERRLTRRHEDGEFWSFSDFLRHYSRLIEICNLTPDTLTSPTYKKMLTK 360
QY 361 MDGNWRGSAAGCGRNTPNTFMNPOYLKLEEBDEDEBEGSGCTFLVGLIQHRRROR 420
DB 361 MDGNWRGSAAGCGRNTPNTFMNPOYLKLEEBDEDEBEGSGCTFLVGLIQHRRROR 420
QY 421 KMGEDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNARERSDPTFLNREVLNRFKLPBG 480
DB 421 KMGEDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNARERSDPTFLNREVLNRFKLPBG 480
QY 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVNDEIEANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVNDEIEANLEEFDISDDIDGVRRLFAQ 540
QY 541 LAGSDAISAPELQTIIRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGIKEFYIL 600
DB 541 LAGSDAISAPELQTIIRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGIKEFYIL 600
QY 601 WTIKIQOKIYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
DB 601 WTIKIQOKIYRELDVDRSGTMSYEMRKALBEAGFMPCCQLHOVIYARFADQOLIIFD 660
QY 661 NFVRCIVRLLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCIVRLLETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

```

RESULT 5
US-09-562-979-9
; Sequence 9, Application US/09562979
; GENERAL INFORMATION:
; APPLICANT: Fredelking, Terry M.
; APPLICANT: Ignatyev, George M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS

FILE REFERENCE: 24881-301B
CURRENT APPLICATION NUMBER: US/09/562.979
CURRENT FILING DATE: 2000-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-09-562-979-9

Query Match 100.0%; Score 3713; DB 19; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLADREABEGSHERAIKYINODYEAALNECEAGTLPDPSPFPAIPALGF 60
DB 1 MAGIAAKLADREABEGSHERAIKYINODYEAALNECEAGTLPDPSPFPAIPALGF 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCMLAAIASLTINEI 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCMLAAIASLTINEI 120
QY 121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
QY 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALOKSLLGC 240
DB 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALOKSLLGC 240
QY 241 SIDTSAADSEALTFQKLVKHAAYVTGAEEVENSGLQKIRIRNPMGEVENTGRWND 300
DB 241 SIDTSAADSEALTFQKLVKHAAYVTGAEEVENSGLQKIRIRNPMGEVENTGRWND 300
QY 301 CPSWNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKWLTJK 360
DB 301 CPSWNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKWLTJK 360
QY 361 MDGMMRRGSTAGGCRNYPNTFMNPOYLKLEEBDEDEDESGCTFLVGLIKHRRRQ 420
DB 361 MDGMMRRGSTAGGCRNYPNTFMNPOYLKLEEBDEDEDESGCTFLVGLIKHRRRQ 420
QY 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARERSDTFINLREVLNRFKLP 480
DB 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARERSDTFINLREVLNRFKLP 480
QY 481 EYIIVPSTFEENKQDGFICIRVFSEKKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
DB 481 EYIIVPSTFEENKQDGFICIRVFSEKKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSIETCKIWMMDLSDSGSLGKLEFYIL 600
DB 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSIETCKIWMMDLSDSGSLGKLEFYIL 600
QY 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHOVIVARFADDOILLIDF 660
DB 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHOVIVARFADDOILLIDF 660
QY 661 NFVRCVLETLFKIKQALDPENTGIIELDISWLCFSVL 700
DB 661 NFVRCVLETLFKIKQALDPENTGIIELDISWLCFSVL 700

Sequence 23, Application US/09768877
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: ODA, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/768.877
CURRENT FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 09/422,869
PRIOR FILING DATE: 1999-10-21
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 700
TYPE: PRT
ORGANISM: Human
US-09-768-877-23

Query Match 100.0%; Score 3713; DB 22; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLADREABEGSHERAIKYINODYEAALNECEAGTLPDPSPFPAIPALGF 60
DB 1 MAGIAAKLADREABEGSHERAIKYINODYEAALNECEAGTLPDPSPFPAIPALGF 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCMLAAIASLTINEI 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALGDCMLAAIASLTINEI 120
QY 121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNQSFOENYAGIFHFQFQWYGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
QY 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALOKSLLGC 240
DB 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALOKSLLGC 240
QY 241 SIDTSAADSEALTFQKLVKHAAYVTGAEEVENSGLQKIRIRNPMGEVENTGRWND 300
DB 241 SIDTSAADSEALTFQKLVKHAAYVTGAEEVENSGLQKIRIRNPMGEVENTGRWND 300
QY 301 CPSWNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKWLTJK 360
DB 301 CPSWNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKWLTJK 360
QY 361 MDGMMRRGSTAGGCRNYPNTFMNPOYLKLEEBDEDEDESGCTFLVGLIKHRRRQ 420
DB 361 MDGMMRRGSTAGGCRNYPNTFMNPOYLKLEEBDEDEDESGCTFLVGLIKHRRRQ 420
QY 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARERSDTFINLREVLNRFKLP 480
DB 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARERSDTFINLREVLNRFKLP 480
QY 481 EYIIVPSTFEENKQDGFICIRVFSEKKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
DB 481 EYIIVPSTFEENKQDGFICIRVFSEKKADYQAVDEIEANLEBPDISEDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSIETCKIWMMDLSDSGSLGKLEFYIL 600
DB 541 LAGEDAEISAFELQTIIRVLAARODIKSDGFSIETCKIWMMDLSDSGSLGKLEFYIL 600
QY 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHOVIVARFADDOILLIDF 660
DB 601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALEAGFMPCOLHOVIVARFADDOILLIDF 660

QY 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

RESULT 7

US-09-791-537-81010
; Sequence 81010, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 81010
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-81010

Query Match 100.0%; Score 3713; DB 22; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDRBAAGLSHERAIKYLNODYALNECLEAGTLFODSPFPAISALGF 60
DB 1 MAGIAKLAKDRBAAGLSHERAIKYLNODYALNECLEAGTLFODSPFPAISALGF 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLAIAISLTINEE 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLAIAISLTINEE 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWYELKKRPPLFKIIOKALOKGSLIGC 240
DB 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWYELKKRPPLFKIIOKALOKGSLIGC 240
QY 241 SIDITSAADSEAITPQKLVKGHAYSVTGAEEVENSNGSLQKLRIRNPMGSEVETGRMND 300
DB 241 SIDITSAADSEAITPQKLVKGHAYSVTGAEEVENSNGSLQKLRIRNPMGSEVETGRMND 300
QY 301 CPSEMTIDPEERERLTRRHEDGEFWMFSFDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
DB 301 CPSEMTIDPEERERLTRRHEDGEFWMFSFDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
QY 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEBDSGCTFLVGLQKHRROR 420
DB 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEBDSGCTFLVGLQKHRROR 420
QY 421 KMGEDHTTGFGLYVPEELSGQTNHLNPNFLTRARERSDPTINLRVLRFPKLPFG 480
DB 421 KMGEDHTTGFGLYVPEELSGQTNHLNPNFLTRARERSDPTINLRVLRFPKLPFG 480
QY 481 EYILVSTEEPNDGFCIRVSEKKADYQAVDEIEANLEEDFIDSDIDGVRRLFAQ 540
DB 481 EYILVSTEEPNDGFCIRVSEKKADYQAVDEIEANLEEDFIDSDIDGVRRLFAQ 540
QY 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIIVMDLSDSGSKLGKEFYIL 600
DB 541 LAGEDAIEISAFELQTLIRVLAARODIKSDGFSIETCKIIVMDLSDSGSKLGKEFYIL 600
QY 601 WTKIOYKOKYRELDVDRSGTMSYEMRKALBEAGFMPCOLHOVIVARADDOQLIIDFD 660
DB 601 WTKIOYKOKYRELDVDRSGTMSYEMRKALBEAGFMPCOLHOVIVARADDOQLIIDFD 660

QY 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCVRLTETLFKIFKQDLPENTGTIELDLISWLCFSVL 700

RESULT 8

US-09-840-707A-9
; Sequence 9, Application US/09840707A
; GENERAL INFORMATION:
; APPLICANT: Fredking, Terry M.
; APPLICANT: Ignatyev, George M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; FILE REFERENCE: 24881-301C
; CURRENT APPLICATION NUMBER: US/09/840,707A
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/562,979
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/198,210
; PRIOR FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: IL-1 receptor intracellular ligand protein
; OTHER INFORMATION: competing amino acid sequence.
; PATENT FILING DATE: 1995-06-07
; PUBLICATION DATE: 1998-10-06
US-09-840-707A-9

Query Match 100.0%; Score 3713; DB 23; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAKDRBAAGLSHERAIKYLNODYALNECLEAGTLFODSPFPAISALGF 60
DB 1 MAGIAKLAKDRBAAGLSHERAIKYLNODYALNECLEAGTLFODSPFPAISALGF 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLAIAISLTINEE 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLAIAISLTINEE 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWYELKKRPPLFKIIOKALOKGSLIGC 240
DB 181 LERAYAKINGCYEALSGATTGEGEDFTGIAEWYELKKRPPLFKIIOKALOKGSLIGC 240
QY 241 SIDITSAADSEAITPQKLVKGHAYSVTGAEEVENSNGSLQKLRIRNPMGSEVETGRMND 300
DB 241 SIDITSAADSEAITPQKLVKGHAYSVTGAEEVENSNGSLQKLRIRNPMGSEVETGRMND 300
QY 301 CPSEMTIDPEERERLTRRHEDGEFWMFSFDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
DB 301 CPSEMTIDPEERERLTRRHEDGEFWMFSFDFLRHYSRLICNLTPTLTSDTYKKWKLTK 360
QY 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEBDSGCTFLVGLQKHRROR 420
DB 361 MDGNMRRGSTAGCCRYNPTFMNPPQYLKLEEDDEDEBDSGCTFLVGLQKHRROR 420
QY 421 KMGEDHTTGFGLYVPEELSGQTNHLNPNFLTRARERSDPTINLRVLRFPKLPFG 480
DB 421 KMGEDHTTGFGLYVPEELSGQTNHLNPNFLTRARERSDPTINLRVLRFPKLPFG 480
QY 481 EYILVSTEEPNDGFCIRVSEKKADYQAVDEIEANLEEDFIDSDIDGVRRLFAQ 540
DB 481 EYILVSTEEPNDGFCIRVSEKKADYQAVDEIEANLEEDFIDSDIDGVRRLFAQ 540

Db 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDPISDDIDDGVRRLFAQ 540
QY 541 LAGEDAISAPELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKGLKEFYIL 600
Db 541 LAGEDAISAPELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKGLKEFYIL 600
QY 601 WTKIQKQKIYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
Db 601 WTKIQKQKIYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
QY 661 NFVRCVLRLETLFKIRKQDPENTGTIELDLISWLCFSYL 700
Db 661 NFVRCVLRLETLFKIRKQDPENTGTIELDLISWLCFSYL 700

RESULT 9

US-09-840-707B-9
; Sequence 9, Application US/09840707B
; GENERAL INFORMATION:
; APPLICANT: Fredeking, Terry M.
; APPLICANT: Ignatyev, George M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
; TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS
; FILE REFERENCE: 24881-301C
; CURRENT APPLICATION NUMBER: US/09/840, 707B
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/562,979
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/198,210
; PRIOR FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: IL-1 receptor intracellular ligand and protein
; OTHER INFORMATION: comprising amino acid sequence
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 5,817,476
; PATENT FILING DATE: 1995-06-07
; PUBLICATION DATE: 1998-10-06
; US-09-840-707B-9

Query Match 100.0%; Score 3713; DB 23; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKREAAEGLSHERAIKYNQVEALRNECLBAGTLFDDSPFPAIPALGF 60
Db 1 MAGIAAKLAKREAAEGLSHERAIKYNQVEALRNECLBAGTLFDDSPFPAIPALGF 60
QY 61 KELGPYSKTRGMKRPTEICADPOFIIGATRTDICOAGLGDCLMAIASLTINEI 120
Db 61 KELGPYSKTRGMKRPTEICADPOFIIGATRTDICOAGLGDCLMAIASLTINEI 120
QY 121 LARVVPNOSFOENAGIFHFQFQWYGEWVEVVDRLPTKDGELLFVHSAEGSEFWSAL 180
Db 121 LARVVPNOSFOENAGIFHFQFQWYGEWVEVVDRLPTKDGELLFVHSAEGSEFWSAL 180
QY 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIQKLOKQSLIGC 240
Db 181 LEKAVAKINGCYEALSGATTEGFEDFTGIAEWELKKPPNLFKIQKLOKQSLIGC 240
QY 241 SIDITSADSEALITFQKLVKHAIVTGAEEVNSGSIQKLIIRINPWGEVWETGRMND 300
Db 241 SIDITSADSEALITFQKLVKHAIVTGAEEVNSGSIQKLIIRINPWGEVWETGRMND 300
QY 301 CPSWNTIDPEERERLTRHDEGEFWMSPDLRAHYSRLICNLTPDTLTSDTYKKMKLT 360
Db 301 CPSWNTIDPEERERLTRHDEGEFWMSPDLRAHYSRLICNLTPDTLTSDTYKKMKLT 360

QY 361 MDGNMRRGSTAGCCRNYPNTFMNNPOYLKLEBEDDEEDGSGCTFLVGLIQKHRROR 420
Db 361 MDGNMRRGSTAGCCRNYPNTFMNNPOYLKLEBEDDEEDGSGCTFLVGLIQKHRROR 420
QY 421 KMEDNHTTGFQIYEPPEELSGOTNHLKINFLTRARERSTFTNLAEVNRFFLP 480
Db 421 KMEDNHTTGFQIYEPPEELSGOTNHLKINFLTRARERSTFTNLAEVNRFFLP 480
QY 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDPISDDIDDGVRRLFAQ 540
Db 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDPISDDIDDGVRRLFAQ 540
QY 541 LAGEDAISAPELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKGLKEFYIL 600
Db 541 LAGEDAISAPELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKGLKEFYIL 600
QY 601 WTKIQKQKIYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
Db 601 WTKIQKQKIYREIDVDRSGTMSYEMRKALAEAGFKMPCQLHQVIVARFADQLIIDPD 660
QY 661 NFVRCVLRLETLFKIRKQDPENTGTIELDLISWLCFSYL 700
Db 661 NFVRCVLRLETLFKIRKQDPENTGTIELDLISWLCFSYL 700

RESULT 10

US-09-884-319-7
; Sequence 7, Application US/09884319
; GENERAL INFORMATION:
; APPLICANT: Lin, Lih-Ling
; APPLICANT: Graham, James
; TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
; INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
; BINDING

NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/884,319
FILING DATE: 18-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,516
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/487,942
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-884-319-7

Query Match 100.0%; Score 3713; DB 23; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGIAKADRAAAGLGSHERAIKYNODYALNNECEAGTLFQDSFPALPSALGF 60
 1 MAGIAKADRAAAGLGSHERAIKYNODYALNNECEAGTLFQDSFPALPSALGF 60
 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWMLAIASTLTNEE 120
 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWMLAIASTLTNEE 120
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKQELLFVHSAEGSEFMSAL 180
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKQELLFVHSAEGSEFMSAL 180
 181 LEKAYAKINGCYALSGGATTBEGEDFTGIAEWYELKKPPNLFKIIQKALOKSILGC 240
 181 LEKAYAKINGCYALSGGATTBEGEDFTGIAEWYELKKPPNLFKIIQKALOKSILGC 240
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEEVSNLSLOKLRIRNPMGEVETGRWND 300
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEEVSNLSLOKLRIRNPMGEVETGRWND 300
 301 CPSEMTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 301 CPSEMTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 361 MDGNMRGSGTAGCCRNYPNTFWMNPOLYLLKEEDEDDEBEGSGCTFLVGLIOKRRROR 420
 361 MDGNMRGSGTAGCCRNYPNTFWMNPOLYLLKEEDEDDEBEGSGCTFLVGLIOKRRROR 420
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLSKNFELTNARARSPTFLNREVLNRFKLP 480
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLSKNFELTNARARSPTFLNREVLNRFKLP 480
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDELEANLEBDESDIDDCVRRLLFAQ 540
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDELEANLEBDESDIDDCVRRLLFAQ 540
 541 LAGEADAISAFELQTLIRRLARODIKSDGFSIETCKIWMVMDLSDSGSKLGKEFYIL 600
 541 LAGEADAISAFELQTLIRRLARODIKSDGFSIETCKIWMVMDLSDSGSKLGKEFYIL 600
 601 WTQIOYOKIYREIDVDSGTMSYEMRKALEAGFMPCOLHOVIVARADOLITD 660
 601 WTQIOYOKIYREIDVDSGTMSYEMRKALEAGFMPCOLHOVIVARADOLITD 660
 661 NFVRCIVRLLETLFKIFKQLDPEMTGTELDLISMLCFSVL 700
 661 NFVRCIVRLLETLFKIFKQLDPEMTGTELDLISMLCFSVL 700

RESULT 11

US-09-884-319A-7

Sequence 7, Application US/09884319A

GENERAL INFORMATION:

APPLICANT: Lit, Lih-ling

INTRA-CELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND

BINDING

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: USA

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/884,319A

FILING DATE: 18-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/083,516

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brown, Scott A.

REGISTRATION NUMBER: 32,724

REFERENCE/DOCKET NUMBER: G15258

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 700 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-884-319A-7

Query Match 100.0%; Score 3713; DB 23; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAGIAKADRAAAGLGSHERAIKYNODYALNNECEAGTLFQDSFPALPSALGF 60
 1 MAGIAKADRAAAGLGSHERAIKYNODYALNNECEAGTLFQDSFPALPSALGF 60
 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWMLAIASTLTNEE 120
 61 KELGPSYSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWMLAIASTLTNEE 120
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKQELLFVHSAEGSEFMSAL 180
 121 LARVPLNOSFOENYAGIFHFQFQWYGEWEVVVDRLPTKQELLFVHSAEGSEFMSAL 180
 181 LEKAYAKINGCYALSGGATTBEGEDFTGIAEWYELKKPPNLFKIIQKALOKSILGC 240
 181 LEKAYAKINGCYALSGGATTBEGEDFTGIAEWYELKKPPNLFKIIQKALOKSILGC 240
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEEVSNLSLOKLRIRNPMGEVETGRWND 300
 241 SIDITSAADSEALTFOKLVKGHAYSVTGAEEVSNLSLOKLRIRNPMGEVETGRWND 300
 301 CPSEMTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 301 CPSEMTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLICNLTPDTLSDTYKKWLTJK 360
 361 MDGNMRGSGTAGCCRNYPNTFWMNPOLYLLKEEDEDDEBEGSGCTFLVGLIOKRRROR 420
 361 MDGNMRGSGTAGCCRNYPNTFWMNPOLYLLKEEDEDDEBEGSGCTFLVGLIOKRRROR 420
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLSKNFELTNARARSPTFLNREVLNRFKLP 480
 421 KMGEDHTTIGFGIYEVPEELSGQTNHLSKNFELTNARARSPTFLNREVLNRFKLP 480
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDELEANLEBDESDIDDCVRRLLFAQ 540
 481 EYLLVSTEEPNDGFCIRVFSEKADYQAVDELEANLEBDESDIDDCVRRLLFAQ 540
 541 LAGEADAISAFELQTLIRRLARODIKSDGFSIETCKIWMVMDLSDSGSKLGKEFYIL 600
 541 LAGEADAISAFELQTLIRRLARODIKSDGFSIETCKIWMVMDLSDSGSKLGKEFYIL 600

QY 601 WTKIOKQKTYREIDVDSGTMSYEMRKALBEAGFMPQCLHQVIVARFADQLIIDPD 660
DB 601 WTKIOKQKTYREIDVDSGTMSYEMRKALBEAGFMPQCLHQVIVARFADQLIIDPD 660
QY 661 NFVRCIVRLTETLFKIFKQDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCIVRLTETLFKIFKQDPENTGTIELDLISWLCFSVL 700

RESULT 12

US-10-038-557A-9
Sequence 9, Application US/10038557A
GENERAL INFORMATION:
APPLICANT: Fredeking, Terry M.
APPLICANT: Ignatyev, George M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS
FILE REFERENCE: 24881-301D
CURRENT APPLICATION NUMBER: US/10/038, 557A
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 09/840,707
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/562,979
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/198,210
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
OTHER INFORMATION: comprising amino acid sequence
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-10-038-557A-9

Query Match 100.0%; Score 3713; DB 26; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLADREAAELGSGHERAIKYLNDYALNECLEAGTLFQDPSPALPSALGF 60
DB 1 MAGIAAKLADREAAELGSGHERAIKYLNDYALNECLEAGTLFQDPSPALPSALGF 60
QY 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAAIASLTINEEI 120
DB 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAAIASLTINEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVVDRLPTKDGELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVVDRLPTKDGELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGGATTEGFEEDFTGGIAEWEYLKRPPLFKIIQKALQKGSILGC 240
DB 181 LERAYAKINGCYEALSGGATTEGFEEDFTGGIAEWEYLKRPPLFKIIQKALQKGSILGC 240
QY 241 SIDITSAADSEATTFQCLVKGHAYSVTGAEEVESNGSLQKILIRNPMGEVEMTGRWNN 300
DB 241 SIDITSAADSEATTFQCLVKGHAYSVTGAEEVESNGSLQKILIRNPMGEVEMTGRWNN 300
QY 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360
QY 361 MDGWMRGSTAGGCRNYPNTFMNPOYLKLEEDDEBDESGCTPLVGLIQGRRROR 420
DB 361 MDGWMRGSTAGGCRNYPNTFMNPOYLKLEEDDEBDESGCTPLVGLIQGRRROR 420

QY 421 KMGEDNHTIGFQIYEVPEELSGQTNILSKNPELJINRAERSDTFNLREVLNRFKLPQ 480
DB 421 KMGEDNHTIGFQIYEVPEELSGQTNILSKNPELJINRAERSDTFNLREVLNRFKLPQ 480
QY 481 EYTLVSTPEPNKDGFCIRVSEKADYQAVNDEIANLEEDFIEDDIDGVRRLFAQ 540
DB 481 EYTLVSTPEPNKDGFCIRVSEKADYQAVNDEIANLEEDFIEDDIDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTLIRRYLAKODIKSGFSIETCKIMVDLDSGSKGLKEFYIL 600
DB 541 LAGEDAEISAFELQTLIRRYLAKODIKSGFSIETCKIMVDLDSGSKGLKEFYIL 600
QY 601 WTKIOKQKTYREIDVDSGTMSYEMRKALBEAGFMPQCLHQVIVARFADQLIIDPD 660
DB 601 WTKIOKQKTYREIDVDSGTMSYEMRKALBEAGFMPQCLHQVIVARFADQLIIDPD 660
QY 661 NFVRCIVRLTETLFKIFKQDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCIVRLTETLFKIFKQDPENTGTIELDLISWLCFSVL 700

RESULT 13

US-10-116-519-12
Sequence 12, Application US/10116519
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
TITLE OF INVENTION: SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
FILE REFERENCE: D0124 NP
CURRENT APPLICATION NUMBER: US/10/116, 519
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/281,253
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/288,768
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/296,180
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/300,620
PRIOR FILING DATE: 2001-06-25
NUMBER OF SEQ ID NOS: 145
SOFTWARE: PatentIn version 3.0
SEQ ID NO 12
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
US-10-116-519-12

Query Match 100.0%; Score 3713; DB 27; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAAKLADREAAELGSGHERAIKYLNDYALNECLEAGTLFQDPSPALPSALGF 60
DB 1 MAGIAAKLADREAAELGSGHERAIKYLNDYALNECLEAGTLFQDPSPALPSALGF 60
QY 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAAIASLTINEEI 120
DB 61 KEIGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAAIASLTINEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVVDRLPTKDGELLFVHSAEGSEFWSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQWQYGEWVEVVVDRLPTKDGELLFVHSAEGSEFWSAL 180
QY 181 LERAYAKINGCYEALSGGATTEGFEEDFTGGIAEWEYLKRPPLFKIIQKALQKGSILGC 240
DB 181 LERAYAKINGCYEALSGGATTEGFEEDFTGGIAEWEYLKRPPLFKIIQKALQKGSILGC 240
QY 241 SIDITSAADSEATTFQCLVKGHAYSVTGAEEVESNGSLQKILIRNPMGEVEMTGRWNN 300
DB 241 SIDITSAADSEATTFQCLVKGHAYSVTGAEEVESNGSLQKILIRNPMGEVEMTGRWNN 300
QY 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360
DB 301 CPSMNTIDPEERERLTRRHEDGEFWSFDFLRHYSRLICNLTPDITLSDTYKKWKLTK 360

Db 301 CPSMNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLCMTLPDTLTSDTYKKWLTk 360
Qy 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
Db 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
Qy 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Db 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Qy 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDDIDGVRRLFAQ 540
Db 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDDIDGVRRLFAQ 540
Qy 541 LAGEDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
Db 541 LAGEDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
Qy 601 WTKIOYKQIYREIDVDRSGTMSYEMRKALBEAGFMPQOLHGVIVARPADQOLIIFD 660
Db 601 WTKIOYKQIYREIDVDRSGTMSYEMRKALBEAGFMPQOLHGVIVARPADQOLIIFD 660
Qy 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700
Db 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700

RESULT 14

US-10-219-051B-12485
; Sequence 12485, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 12485
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Refseq / NP_001739
; DATABASE ENTRY DATE: 2001-01-23
US-10-219-051B-12485

Query Match 100.0%; Score 3713; DB 28; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIAKLADREARAGLSHERAIKYLNDYEALNECELEAGTLFQDPSFPALPSALGp 60
Db 1 MAGIAKLADREARAGLSHERAIKYLNDYEALNECELEAGTLFQDPSFPALPSALGp 60
Qy 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAIAISLTINEE 120
Db 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAIAISLTINEE 120
Qy 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAESSEFWSAL 180
Db 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAESSEFWSAL 180
Qy 181 LERAYAKINCYEALSGATTEGFEPTGGIAEWYELKRPPLFKIIOALOKSGLGC 240
Db 181 LERAYAKINCYEALSGATTEGFEPTGGIAEWYELKRPPLFKIIOALOKSGLGC 240

Db 181 LERAYAKINCYEALSGATTEGFEPTGGIAEWYELKRPPLFKIIOALOKSGLGC 240
Qy 241 SIDTSAADSEAITFOKLKYGAYSVTGAEEVENSGLQKILIRNPMGEVETGRMND 300
Db 241 SIDTSAADSEAITFOKLKYGAYSVTGAEEVENSGLQKILIRNPMGEVETGRMND 300
Qy 301 CPSMNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLCMTLPDTLTSDTYKKWLTk 360
Db 301 CPSMNTIDPEERERLTRRHEDGEFWMSPDFLRHYSRLCMTLPDTLTSDTYKKWLTk 360
Qy 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
Db 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLLKLEEDDEDEDESGCTFLVGLIOKHRROR 420
Qy 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Db 421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNBARERSPTFNLREVLNRFKLPg 480
Qy 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDDIDGVRRLFAQ 540
Db 481 EYLLVSTPEPNKDGFCIRVFSEKADYQAVDEIEANLEEDDIDGVRRLFAQ 540
Qy 541 LAGEDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
Db 541 LAGEDAEISAFELQTLIRRLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
Qy 601 WTKIOYKQIYREIDVDRSGTMSYEMRKALBEAGFMPQOLHGVIVARPADQOLIIFD 660
Db 601 WTKIOYKQIYREIDVDRSGTMSYEMRKALBEAGFMPQOLHGVIVARPADQOLIIFD 660
Qy 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700
Db 661 NFVRCVLRLETLFKIFKQOLDPENTGTIELDLISWLCFSYL 700

RESULT 15

US-10-293-017-78
; Sequence 78, Application US/10293017
; GENERAL INFORMATION:
; APPLICANT: Shamloo, Mehrdad
; APPLICANT: Chin, Daniel J.
; APPLICANT: Von Schack, David
; APPLICANT: Ufer, Roman
; APPLICANT: Patchy, Laszlo
; APPLICANT: Wieloch, Tadeusz
; APPLICANT: Gonzalez-Zulueta, Mirella
; TITLE OF INVENTION: Regulated Genes in the Pathophysiology of Ischemic Stroke
; FILE REFERENCE: AGYT-007U51
; CURRENT APPLICATION NUMBER: US/10/293,017
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 60/384,522
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/338,296
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 78
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-017-78

Query Match 100.0%; Score 3713; DB 28; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGIAKLADREARAGLSHERAIKYLNDYEALNECELEAGTLFQDPSFPALPSALGp 60
Db 1 MAGIAKLADREARAGLSHERAIKYLNDYEALNECELEAGTLFQDPSFPALPSALGp 60
Qy 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAIAISLTINEE 120
Db 61 KELGPISSKTRGMRKRPTEICADPOFIIGATRTDICOGALGDCWLLAIAISLTINEE 120

Qy 121 LARVPLNOSFOENYAGIFHFQFMQYGEWVEVVDDRLPTKGGELLFVHSAEGSEFMSAL 180
| | | | |
Db 121 LARVPLNOSFOENYAGIFHFQFMQYGEWVEVVDDRLPTKGGELLFVHSAEGSEFMSAL 180
| | | | |
Qy 181 LEKAYAKINGCYEALSGATTEGFEFTGGIAEWYELKKPPNLFKIIQKALQKSLG 240
| | | | |
Db 181 LEKAYAKINGCYEALSGATTEGFEFTGGIAEWYELKKPPNLFKIIQKALQKSLG 240
| | | | |
Qy 241 SIDTSAADSALTFQKLVKGHAYSVTGAEEVESNGSLQKLRIRNPGGEVEMTGRMND 300
| | | | |
Db 241 SIDTSAADSALTFQKLVKGHAYSVTGAEEVESNGSLQKLRIRNPGGEVEMTGRMND 300
| | | | |
Qy 301 CPSWNTIDPEERELTRRHEDGEPFMSFDELRAHYSRLICNLTPDILTSPTYKKMKLTK 360
| | | | |
Db 301 CPSWNTIDPEERELTRRHEDGEPFMSFDELRAHYSRLICNLTPDILTSPTYKKMKLTK 360
| | | | |
Qy 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
| | | | |
Db 361 MDGNMRGSGTAGGCRNYPNTFWMNPQYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
| | | | |
Qy 421 KMGDMHTIGFGIYEVPEELSGQTNHLSKNFPLTNRARESDTFINLREVLNRFKLP 480
| | | | |
Db 421 KMGDMHTIGFGIYEVPEELSGQTNHLSKNFPLTNRARESDTFINLREVLNRFKLP 480
| | | | |
Qy 481 EYILVPSTFEPNKDGDFCIRVFSEKADYQAVDEIEANLEEDISEDDIDDGVRLLFAQ 540
| | | | |
Db 481 EYILVPSTFEPNKDGDFCIRVFSEKADYQAVDEIEANLEEDISEDDIDDGVRLLFAQ 540
| | | | |
Qy 541 LAGEDAEISAFELQTIIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGJKEFYIL 600
| | | | |
Db 541 LAGEDAEISAFELQTIIRVLAKRQDIKSDGFSIETCKIMVMDLSDSGKLGJKEFYIL 600
| | | | |
Qy 601 WTKIQXOKIYREIDVDSGTMSYEMRKALEEAGFKPCQLHOYIVARFADQLIIDFD 660
| | | | |
Db 601 WTKIQXOKIYREIDVDSGTMSYEMRKALEEAGFKPCQLHOYIVARFADQLIIDFD 660
| | | | |
Qy 661 NFVACLVRLETLFKIFKQIDPENTGTELDLISWLCFSVL 700
| | | | |
Db 661 NFVACLVRLETLFKIFKQIDPENTGTELDLISWLCFSVL 700
| | | | |

Search completed: July 24, 2003, 13:12:46
Job time : 424.806 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:53:55 ; Search time 47.4084 Seconds
(without alignments)
1753.524 Million cell updates/sec

Title: US-09-884-319a-7
Perfect score: 3713
Sequence: 1 MAGIAKLADREAEGLSGS.....PENTGTELDLISWLCPSVL 700

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3713	100.0	700	US-09-840-707a-9	Sequence 9, Appli
2	3713	100.0	700	US-09-768-877-23	Sequence 23, Appli
3	3713	100.0	700	US-09-884-319-7	Sequence 7, Appli
4	3713	100.0	700	US-10-038-557a-9	Sequence 9, Appli
5	3713	100.0	700	US-10-116-519-12	Sequence 12, Appli
6	2458.5	66.2	714	US-09-768-877-22	Sequence 22, Appli
7	2458.5	66.2	714	US-10-116-519-6	Sequence 6, Appli
8	2435	65.6	713	US-10-116-519-8	Sequence 8, Appli
9	2356.5	63.5	702	US-09-768-877-27	Sequence 27, Appli
10	2014	54.2	703	US-10-116-519-11	Sequence 11, Appli
11	1943.5	52.3	709	US-10-116-519-9	Sequence 9, Appli
12	1882	50.7	690	US-09-768-877-28	Sequence 28, Appli
13	1882	50.7	690	US-10-116-519-13	Sequence 13, Appli
14	1862.5	50.2	821	US-09-768-877-24	Sequence 24, Appli
15	1862.5	50.2	821	US-10-116-519-7	Sequence 7, Appli

16	1795	48.3	664	US-10-325-881-70	Sequence 70, Appli
17	1783.5	48.0	647	US-10-116-519-10	Sequence 10, Appli
18	1287	34.7	251	US-09-840-707a-8	Sequence 8, Appli
19	1287	34.7	251	US-09-884-319-6	Sequence 6, Appli
20	1287	34.7	251	US-10-038-557a-8	Sequence 8, Appli
21	1150	31.0	697	US-10-116-519-56	Sequence 56, Appli
22	1148.5	30.9	694	US-10-116-519-54	Sequence 54, Appli
23	1126.5	30.3	343	US-09-794-960-4	Sequence 221, App
24	1056	28.4	206	US-10-043-487-221	Sequence 5, Appli
25	1020	27.5	300	US-09-794-960-5	Sequence 2, Appli
26	938	25.3	428	US-10-116-519-2	Sequence 25, Appli
27	938	25.3	428	US-10-116-519-25	Sequence 25, Appli
28	901.5	24.3	639	US-09-768-877-25	Sequence 2, Appli
29	893	24.1	639	US-10-116-519-5	Sequence 2, Appli
30	845.5	22.8	669	US-09-962-739-2	Sequence 26, Appli
31	780	21.0	641	US-09-768-877-26	Sequence 4, Appli
32	656	17.7	353	US-09-962-739-4	Sequence 92, Appli
33	635	17.1	242	US-10-116-519-92	Sequence 93, Appli
34	635	17.1	242	US-10-116-519-93	Sequence 46, Appli
35	632.5	17.0	241	US-10-116-519-46	Sequence 18, Appli
36	621.5	16.7	666	US-09-768-877-18	Sequence 4, Appli
37	607	16.3	544	US-09-768-877-4	Sequence 4, Appli
38	604	16.3	671	US-10-116-519-4	Sequence 2, Appli
39	604	16.3	672	US-09-768-877-2	Sequence 2, Appli
40	600	16.2	1603	US-10-208-948-2	Sequence 24, Appli
41	600	16.2	2159	US-10-208-948-24	Sequence 8, Appli
42	599	16.1	513	US-09-768-877-8	Sequence 2, Appli
43	579	15.6	517	US-09-794-960-2	Sequence 6, Appli
44	575	15.5	517	US-09-768-877-6	Sequence 27, Appli
45	561	15.1	2150	US-10-208-948-27	

ALIGNMENTS

RESULT 1
US-09-840-707a-9
Sequence 9, Application US/09840707A
Patent No. US20020077276A1
GENERAL INFORMATION:
APPLICANT: Predakiev, Terry M.
INVENTOR: Ignatyev, George M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
FILE REFERENCE: 24881-301C
CURRENT APPLICATION NUMBER: US/09/840,707A
CURRENT FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/562,979
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/198,210
PRIOR FILING DATE: 1999-04-27
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 700
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: IL-1 receptor intracellular ligand protein
PUBLICATION INFORMATION:
PATENT DOCUMENT NUMBER: 5,817,476
PATENT FILING DATE: 1995-06-07
PUBLICATION DATE: 1998-10-06
US-09-840-707a-9

Query Match 100.0%; Score 3713; DB 9; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 1 MAGIAKLADREAEGLSGHERAIKYLNDYALRNECEAGTLFDDPSFPAISALGF 60
DB 1 MAGIAKLADREAEGLSGHERAIKYLNDYALRNECEAGTLFDDPSFPAISALGF 60

QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTNEI 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTNEI 120
QY 121 LARVPLNQSFOENYAGIFHFQFQWGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNQSFOENYAGIFHFQFQWGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
QY 181 LERAVAKINGCYEALSGGATTEGFEFTGIAEMWELKPPNLFKIIQKALQKGSILGC 240
DB 181 LERAVAKINGCYEALSGGATTEGFEFTGIAEMWELKPPNLFKIIQKALQKGSILGC 240
QY 241 SIDTSADESEALTFOKLVGHAYSVTGAEEVENSGLQKLIIRNPGVEWMTGRMND 300
DB 241 SIDTSADESEALTFOKLVGHAYSVTGAEEVENSGLQKLIIRNPGVEWMTGRMND 300
QY 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLIECNLTPTDITSDTYKKWLTk 360
DB 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLIECNLTPTDITSDTYKKWLTk 360
QY 361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
DB 361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
QY 421 KMGEDMTTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPg 480
DB 421 KMGEDMTTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPg 480
QY 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVDDEIEANLEEPDISDDIDDGVRRLFAQ 540
DB 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVDDEIEANLEEPDISDDIDDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
QY 601 WTKIQKQKXIREIDVDRSGTMNSYEMRKALBEAGFMPCOLHQVYARFADQOLIIFD 660
DB 601 WTKIQKQKXIREIDVDRSGTMNSYEMRKALBEAGFMPCOLHQVYARFADQOLIIFD 660
QY 661 NFVRCLVRLTFLFKIFKQIDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCLVRLTFLFKIFKQIDPENTGTIELDLISWLCFSVL 700

RESULT 2
US-09-768-877-23
Sequence 23, Application US/09768877
Patent No. US20020150896A1
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOMI
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/768, 877
CURRENT FILING DATE: 2001-01-23
PRIORITY APPLICATION NUMBER: 09/422, 869
PRIORITY FILING DATE: 1999-10-21
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 700
TYPE: PRT
ORGANISM: Human
US-09-768-877-23

Query Match 100.0%; Score 3713; DB 10; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGIAAALADREABEGSHERAIKYNODYALRNECELAGTLPDPSFPAPISALGF 60
DB 1 MAGIAAALADREABEGSHERAIKYNODYALRNECELAGTLPDPSFPAPISALGF 60
QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTNEI 120
DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIASTLTNEI 120
QY 121 LARVPLNQSFOENYAGIFHFQFQWGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNQSFOENYAGIFHFQFQWGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
QY 181 LERAVAKINGCYEALSGGATTEGFEFTGIAEMWELKPPNLFKIIQKALQKGSILGC 240
DB 181 LERAVAKINGCYEALSGGATTEGFEFTGIAEMWELKPPNLFKIIQKALQKGSILGC 240
QY 241 SIDTSADESEALTFOKLVGHAYSVTGAEEVENSGLQKLIIRNPGVEWMTGRMND 300
DB 241 SIDTSADESEALTFOKLVGHAYSVTGAEEVENSGLQKLIIRNPGVEWMTGRMND 300
QY 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLIECNLTPTDITSDTYKKWLTk 360
DB 301 CPSWNTIDPEERERLTRRHEDGEFMSFSDFLRHYSRLIECNLTPTDITSDTYKKWLTk 360
QY 361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
DB 361 MDGMRRGSTAGGCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIOKRRROR 420
QY 421 KMGEDMTTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPg 480
DB 421 KMGEDMTTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPg 480
QY 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVDDEIEANLEEPDISDDIDDGVRRLFAQ 540
DB 481 EYIIVPSTFEPNKGDFCIRVSEKADYQAVDDEIEANLEEPDISDDIDDGVRRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKGLKEFYIL 600
QY 601 WTKIQKQKXIREIDVDRSGTMNSYEMRKALBEAGFMPCOLHQVYARFADQOLIIFD 660
DB 601 WTKIQKQKXIREIDVDRSGTMNSYEMRKALBEAGFMPCOLHQVYARFADQOLIIFD 660
QY 661 NFVRCLVRLTFLFKIFKQIDPENTGTIELDLISWLCFSVL 700
DB 661 NFVRCLVRLTFLFKIFKQIDPENTGTIELDLISWLCFSVL 700

RESULT 3
US-09-884-319-7
Sequence 7, Application US/09884319
Publication No. US20030124625A1
GENERAL INFORMATION:
APPLICANT: Lin, Lih-ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
BINDING
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/884,319
 FILING DATE: 18-Jun-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,516
 FILING DATE: <Unknown>
 APPLICATION NUMBER: 08/487,942
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Brown, Scott A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: G15258
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 700 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 US-09-884-319-7

Query Match 100.0%; Score 3713; DB 11; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAADREABAGLSHERAIKYNODYEALNECELEAGTLFODPSFPAIPSAIGF 60
 DB 1 MAGIAKLAADREABAGLSHERAIKYNODYEALNECELEAGTLFODPSFPAIPSAIGF 60
 QY 61 KELGPYSSKTRGRMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISLTINEE 120
 DB 61 KELGPYSSKTRGRMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISLTINEE 120
 QY 121 LARVPLNOSFOENYAGIFHFQFOWYGEWEVVVDRLPTKDELLFVHSAEGSEFWSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFOWYGEWEVVVDRLPTKDELLFVHSAEGSEFWSAL 180
 QY 121 LARVPLNOSFOENYAGIFHFQFOWYGEWEVVVDRLPTKDELLFVHSAEGSEFWSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFOWYGEWEVVVDRLPTKDELLFVHSAEGSEFWSAL 180
 QY 181 LERAYAKINGCYALSGGATTEGFEDEFTGGIAEWEYLKPPNLFKIIQKALOKGSLGC 240
 DB 181 LERAYAKINGCYALSGGATTEGFEDEFTGGIAEWEYLKPPNLFKIIQKALOKGSLGC 240
 QY 241 SIDITSAADSEALTFOKLYKGHAYSVTGAEEVESNSLOKLIIRINPWGEVETGRWNN 300
 DB 241 SIDITSAADSEALTFOKLYKGHAYSVTGAEEVESNSLOKLIIRINPWGEVETGRWNN 300
 QY 301 CPSEMTIDPEERERLRRHDEGEFWSFSDFLHYSLKCNLTPTDLSDTYKKWKLTK 360
 DB 301 CPSEMTIDPEERERLRRHDEGEFWSFSDFLHYSLKCNLTPTDLSDTYKKWKLTK 360
 QY 361 MDGNMRRGSTAGCCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIQKHRRROR 420
 DB 361 MDGNMRRGSTAGCCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIQKHRRROR 420
 QY 421 KMGEDEHTTIFGFIYEVEELSGOTNHLKSNFPLTNARERSDTFINLRVLRPKLPG 480
 DB 421 KMGEDEHTTIFGFIYEVEELSGOTNHLKSNFPLTNARERSDTFINLRVLRPKLPG 480
 QY 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEFDISIEDDDGVRRLFAQ 540
 DB 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEFDISIEDDDGVRRLFAQ 540
 QY 541 LAEDDAISAFELQTIIRVLAQRDIIKSGFSIETTKIIVMDLDSGSGKLGIKEFYIL 600
 DB 541 LAEDDAISAFELQTIIRVLAQRDIIKSGFSIETTKIIVMDLDSGSGKLGIKEFYIL 600

DB 541 LAEDDAISAFELQTIIRVLAQRDIIKSGFSIETTKIIVMDLDSGSGKLGIKEFYIL 600
 QY 601 WTKIQYKQIKYREIDVDSGTMSYEMRKALAEAGFMPQOLHOVIARFADQIIFD 660
 DB 601 WTKIQYKQIKYREIDVDSGTMSYEMRKALAEAGFMPQOLHOVIARFADQIIFD 660
 QY 661 NFVRCVLRLETLFKIFKQDPENTGTIELDLSWLCFSVL 700
 DB 661 NFVRCVLRLETLFKIFKQDPENTGTIELDLSWLCFSVL 700

RESULT 4
 US-10-038-557A-9
 Sequence 9, Application US/10038557A
 Publication No. US20030092684A1
 GENERAL INFORMATION:
 APPLICANT: Fredeking, Terry M.
 APPLICANT: Ignatyev, George M.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING HEMORRHAGIC VIRUS
 TITLE OF INVENTION: INFECTIONS AND OTHER DISORDERS
 FILE REFERENCE: 24881-301D
 CURRENT APPLICATION NUMBER: US/10/038,557A
 CURRENT FILING DATE: 2002-06-05
 PRIOR APPLICATION NUMBER: 09/840,707
 PRIOR FILING DATE: 2001-04-23
 PRIOR APPLICATION NUMBER: 09/562,979
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: 60/198,210
 PRIOR FILING DATE: 1999-04-27
 NUMBER OF SEQ ID NOS: 26
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 9
 LENGTH: 700
 TYPE: PRT

ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: IL-1 receptor intracellular ligand protein
 OTHER INFORMATION: comprising amino acid sequence
 PUBLICATION INFORMATION:
 PATENT DOCUMENT NUMBER: 5,817,476
 PATENT FILING DATE: 1995-06-07
 PUBLICATION DATE: 1998-10-06
 US-10-038-557A-9

Query Match 100.0%; Score 3713; DB 15; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLAADREABAGLSHERAIKYNODYEALNECELEAGTLFODPSFPAIPSAIGF 60
 DB 1 MAGIAKLAADREABAGLSHERAIKYNODYEALNECELEAGTLFODPSFPAIPSAIGF 60
 QY 61 KELGPYSSKTRGRMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISLTINEE 120
 DB 61 KELGPYSSKTRGRMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISLTINEE 120
 QY 121 LARVPLNOSFOENYAGIFHFQFOWYGEWEVVVDRLPTKDELLFVHSAEGSEFWSAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFOWYGEWEVVVDRLPTKDELLFVHSAEGSEFWSAL 180
 QY 181 LERAYAKINGCYALSGGATTEGFEDEFTGGIAEWEYLKPPNLFKIIQKALOKGSLGC 240
 DB 181 LERAYAKINGCYALSGGATTEGFEDEFTGGIAEWEYLKPPNLFKIIQKALOKGSLGC 240
 QY 241 SIDITSAADSEALTFOKLYKGHAYSVTGAEEVESNSLOKLIIRINPWGEVETGRWNN 300
 DB 241 SIDITSAADSEALTFOKLYKGHAYSVTGAEEVESNSLOKLIIRINPWGEVETGRWNN 300
 QY 301 CPSEMTIDPEERERLRRHDEGEFWSFSDFLHYSLKCNLTPTDLSDTYKKWKLTK 360
 DB 301 CPSEMTIDPEERERLRRHDEGEFWSFSDFLHYSLKCNLTPTDLSDTYKKWKLTK 360
 QY 361 MDGNMRRGSTAGCCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIQKHRRROR 420
 DB 361 MDGNMRRGSTAGCCRNYPNTFMNPOYLKLEEBEDEDESGCTFLVGLIQKHRRROR 420

```

Db      361 MDGMRRGSTAGGCRNYPNTFMNPPQYLKLEEDDEDEDESGCTFLVGLIQKRRROR 420
Qy      421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARSDFINLRVLANFKLPKPG 480
Db      421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARSDFINLRVLANFKLPKPG 480
Qy      481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Db      481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Qy      541 LAGEDAIEISAFELQTLIRRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGLKEFYIL 600
Db      541 LAGEDAIEISAFELQTLIRRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGLKEFYIL 600
Qy      601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPCOLHOVYARFADQOLIIFD 660
Db      601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPCOLHOVYARFADQOLIIFD 660
Qy      661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISMLCFSVL 700
Db      661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISMLCFSVL 700

```

RESULT 5

```

US-10-116-519-12
; Sequence 12, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124 NP
; CURRENT APPLICATION NUMBER: US/10/116,519
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/281,253
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/288,768
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/296,180
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/300,620
; PRIOR FILING DATE: 2001-06-25
; NUMBER OF SEQ ID NOS: 145
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-519-12

```

Query Match 100.0%; Score 3713; DB 15; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MAGIAAKLAKDREAEGLSHERAIKYNODYEALRNECLAGTLFQDPSFPAISALGF 60
Db      1 MAGIAAKLAKDREAEGLSHERAIKYNODYEALRNECLAGTLFQDPSFPAISALGF 60
Qy      61 KELGPYSSKTRGMRWRKPTTEICADPQITIGATRTDICOAGLGDCLAAIASLTINEEI 120
Db      61 KELGPYSSKTRGMRWRKPTTEICADPQITIGATRTDICOAGLGDCLAAIASLTINEEI 120
Qy      121 LARVVLNOSFOENYAGIFHFQFQWYGEWVAVDDRLPTQDGLLFVHSAEGSFPWSAL 180
Db      121 LARVVLNOSFOENYAGIFHFQFQWYGEWVAVDDRLPTQDGLLFVHSAEGSFPWSAL 180
Qy      181 LEKAVAKINGCYEALSGATTEGFDFGIAEWYELKKPPNLFKTIQKLOKGSLLGC 240
Db      181 LEKAVAKINGCYEALSGATTEGFDFGIAEWYELKKPPNLFKTIQKLOKGSLLGC 240
Qy      241 SIDITSADSEAITFOKLKGVHAYSVTGAEVESNGSLQKILIRINPWEVEMTGRWMDN 300

```

```

Db      241 SIDITSADSEAITFOKLKGVHAYSVTGAEVESNGSLQKILIRINPWEVEMTGRWMDN 300
Qy      301 CPSEMTIDPEERERLIRRHEDGEFMMSPDFLHRYGLIEICNLTPDLTSDTYKKWLTK 360
Db      301 CPSEMTIDPEERERLIRRHEDGEFMMSPDFLHRYGLIEICNLTPDLTSDTYKKWLTK 360
Qy      361 MDGMRRGSTAGGCRNYPNTFMNPPQYLKLEEDDEDEDESGCTFLVGLIQKRRROR 420
Db      361 MDGMRRGSTAGGCRNYPNTFMNPPQYLKLEEDDEDEDESGCTFLVGLIQKRRROR 420
Qy      421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARSDFINLRVLANFKLPKPG 480
Db      421 KMGEDMHTIGFGIYEVEPEELSGQTNHLSKNFPLTNRARSDFINLRVLANFKLPKPG 480
Qy      481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Db      481 EYILVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEBEDISEDDIDGVRRLFAQ 540
Qy      541 LAGEDAIEISAFELQTLIRRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGLKEFYIL 600
Db      541 LAGEDAIEISAFELQTLIRRVLAARODIKSDGFSIETCKIWMMLDSGSGKLGLKEFYIL 600
Qy      601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPCOLHOVYARFADQOLIIFD 660
Db      601 WTKIQKQKTYREIDVDRSGTMSYEMRKALAEAGFKMPCOLHOVYARFADQOLIIFD 660
Qy      661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISMLCFSVL 700
Db      661 NFVRCVLRLETLFKIFKQIDPENTGTIELDLISMLCFSVL 700

```

RESULT 6

```

US-09-768-877-22
; Sequence 22, Application US/09768877
; Patent No. US20020150896A1
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/768,877
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/422,869
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 714
; TYPE: PRT
; ORGANISM: Human
US-09-768-877-22

```

Query Match 66.2%; Score 2458.5; DB 10; Length 714;
 Best Local Similarity 63.1%; Pred. No. 6,3e-221;
 Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

```

Qy      3 GIAAKLAKDREAEGLSHERAIKYNODYEALRNECLAGTLFQDPSFPAISALGFKE 62
Db      3 GVSAYQVKQARELIGLRHENAIAIKYIGQDYEQLRVRLOGLTFRDAEPVPOVQSLGYKO 72
Qy      63 LGRPSYSTRGMRWRKPTTEICADPQITIGATRTDICOAGLGDCLAAIASLTINEEILA 122
Db      73 LGRNSSTIYGIKMRPPELLSNPQIYDGAIRTDICOAGLGDCLAAIASLTINDTLH 132
Qy      123 RYVPLNOSFOENYAGIFHFQFQWYGEWVAVDDRLPTQDGLLFVHSAEGSFPWSAL 182

```


Query Match 65.6%; Score 2435; DB 15; Length 713;
Best Local Similarity 62.8%; Pred. No. 1e-218;
Matches 437; Conservative 126; Mismatches 129; Indels 4; Gaps 3;

```
QY 3 GIANKLADREABEAGLSHERAIAKYNODYALNRECEAGTLFODPSPFPAIPSAALGK 62
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 13 GVSAAVOQKREKELGELGHENAIKYLQDYEFLRRACLSQSGVLRQDEAFPPVSHSLGPK 72
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 63 LGPSSKTRGMRKRPTEICADPOPIIGGATRTDICOALDPCWLLAIASTLTNEEIIA 122
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 73 LGPSSKTYGKMKRPTEIMNPOTIVGATRTDICOALDPCWLLAIASTLTNETIILH 132
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 123 RVPLNQSFOENYAGIFHFQFQWQGEWVEVVVDRLPTKDELLFVHSAEGSEFWASLLE 182
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 133 RVFYGOSFOQYAGIFHFQFQWQGEWVDVVIDLLPTKDGKLVFVHSAQNEFEWSALLE 192
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 183 KAVAKINGCYEALSGGATTEGFEPTGIAEWELKRPPLFKIIQALOKSGSLGSI 242
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 193 KAVAKINGCYEALSGGATSEAFEDPTGVTWYDLOKAPSDLYOITLALRSGSLGSI 252
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 243 DITSAADSEALTPOKLVKHAYSVTGAEEVSNGLQKILIRNPMGEVMTGRNDNCP 302
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 253 NISDIRDEALTTPKLVGHAYSTGAKQVYQGRVNLIRKRPWGEVEMKGPMSDSY 312
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 303 SWNTIDPEERBLTRRHEDGFEWMSFSDPLRHYSRLICNTLPDTLTSDTYKKWKLTKMD 362
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 313 EMNKVDPYERBQLRVKMGDEFWMSFRDPIREFTLEICNLTDPALKSRTLIRNMTTTFYE 372
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 363 GNMRRGSTAGGCRNYPNTFMNPOYLKLEBED--EDEDSESGTFLVGLIOKRRRQ 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 373 GTWRRGSTAGGCRNYPATFWNPQKIRLEEVDDADDVNDRESGCSFLALAMQKRRER 432
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 421 KMGDMHTIGIGIYEVPELSGQNTIHLKSNFPLTRARERSDPTINIRVLANRKLTPG 480
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 433 RFGDMETIGIAVYQVPELQGP-VHLKRPFLANRASQSEHFINIRVSNRRLDPPG 491
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 481 EYILVSTFEPNKGDFCIRVFESEKKADYQAVDEIEANL--EEDPISBDIDDDGVRLLFA 539
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 492 EYIVVSTFEPNKGDFLRRFSEKKAQTOELDDIQANLPDEKVLSESEIDDNKTLFS 551
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 540 QLAGEDAIISAFELQTLIRRLVAKRODLSKDSGFSIETKINVMMLDSDSGKLGKEYI 599
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 552 KLAGDMEISVKELOTILNRIISIKKIDRTNGFSLSCRSVNLNDRDGNKGLGIVEPNI 611
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 600 LMTKIOKQKIYREIIVDVSRTMNSYEMRKALBEAGFMPCOLHOVIYARFADDDLI 659
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 612 LMRIRNLTFRKEDDLKSGSMSAYENRMAIEAAGFKLAKKHELIITRYSBPDLAVDF 671
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 660 DNFVRCIVRLTLETKIFKQDLPENTGTIELDLISWL 695
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 672 DNFVCCIVRLTFRFFKLDLTDLDGVTVPDLFKWL 707
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 9
US-09-768-877-27

Sequence 27, Application US/09768877
Patent No. US20020150896A1
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEAMUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG L.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD:307
CURRENT APPLICATION NUMBER: US/09/768,877
CURRENT FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 09/422,869
PRIOR FILING DATE: 1999-10-21

NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 703
TYPE: PRT
ORGANISM: RAT
US-09-768-877-27

Query Match 63.5%; Score 2356.5; DB 10; Length 703;
Best Local Similarity 60.9%; Pred. No. 2.2e-211;
Matches 427; Conservative 127; Mismatches 140; Indels 7; Gaps 3;

```
QY 1 MAGIAKLADREABEAGLSHERAIAKYNODYALNRECEAGTLFODPSPFPAIPSAALG 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 MAALAAVSQKRAVABEAGLSQNAVKYLQDFFELRKQCLNSGVLFPDPEFPACPSALGY 60
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 KEIGPYSKTRGMRKRPTEICADPOPIIGGATRTDICOALDPCWLLAIASTLTNEEII 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 KDLGPGSPDTQGIWKRKPTLCPNPOTIVGATRTDICOALDPCWLLAIASTLTNEKL 120
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 121 LARVPLNQSFOENYAGIFHFQFQWQGEWVEVVVDRLPTKDELLFVHSAEGSEFWASL 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 LRVLPDQSFQDQYAGIFHFQFQWQGEWVEVVIDRLPTKQQLFLHSEBGNFMSAL 180
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 LERAVAKINGCYEALSGGATTEGFEPTGIAEWELKRPPLFKIIQALOKSGSLG 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 LERAVAKINGCYEALVGGSTIEGFEPTGISEFYDLKRPENLYIIOKALRSGSLG 240
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 241 SIDTSAADSEALTPOKLVKHAYSVTGAEEVSNGLQKILIRNPMGEVMTGRNDN 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 SIDTSAAEALTPOKLVGHAYSVTGAEEVSNGLQKILIRNPMGEVMSGAWSDN 300
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 301 CPSWNTIDPEERBLTRRHEDGFEWMSFSDPLRHYSRLICNTLPDTLTSDTYKKWKLTK 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 301 APEWNTIDPEERBLTRRHEDGFEWMSFSDPLQKYSLEICNLTSPDLSSEELHKNLVL 360
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 361 MDGNRRGSTAGGCRNYPNTFMNPOYLKLEBEDDEBDG--ESGCTFLVGLIOKRRR 418
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 FNGWRTGSTAGGCLNPGTYMNPQFKIHLDVEDEQEGTSEPCCTVLLGLMQKRRR 420
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 419 QKKGEMHTIGIGIYEVPELSGQNTIHLKSNFPLTRARERSDPTINIRVLANRKL 478
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 421 QKRGQGMELSIGVAYVQIPLELSHTDAHGROFFLORPSTSSITYMNLREVSSRYRLP 480
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 479 PGXYILVSTFEPNKGDFCIRVFESEKKADYQAVDEIEANL--EEDPISBDIDDDGVR 535
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 481 PGQILVSTFEPNKGDFCIRVFESEKKAKALEIGDVSGHPRHPRDDE--DEHYR 538
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 536 RLFAQLAGEDAISIAPELQTLIRRLVAKRODLSKDSGFSIETKINVMMLDSDSGKLGK 595
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 539 SLFEPFVGKDSSEISANQKRVNLVLSKRTDMKFDGNNINTCREMISLSDSGSLGPM 598
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 596 EYILMTKIOKQKIYREIIVDVSRTMNSYEMRKALBEAGFMPCOLHOVIYARFADDO 655
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 599 EFKTLMKIRTYEIEIFQEMDNHNVGTIEAHMETALKKAPFTLNNVOQTILARVYASKL 658
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 656 IIDDFNFRGVRLTLETKIFKQDLPENTGTIELDLISWL 696
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 659 GVDPNGFVACMIRLETIFKLRLLDKDQNGIVQLSLAEWLJC 699
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
```

RESULT 10
US-10-116-519-11

Sequence 11, Application US/10116519
Publication No. US20030114373A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
FILE REFERENCE: D0124 NP
CURRENT APPLICATION NUMBER: US/10/116,519
CURRENT FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/281,253


```
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/288,768
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/296,180
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/300,620
; PRIOR FILING DATE: 2001-06-25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 702
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-116-519-11

Query Match          54.2%; Score 2014; DB 15; Length 702;
Best Local Similarity 52.4%; Pred. No. 2.4e-179;
Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
```

QY 4 IAAKLADEAAGLSHERAIVKYNODYEALRNECLAGTLFODSPFAIPSAIGFKEIPY 63
1 MVAHINNSRLKAGVGHDNAQNGNSFELRAACLRKELFEDPLFPAEPSSIGFKDL 60
QY 64 GPYSSKTRGWRKRPETICADPOFIIGATRTDICOGLGDCWLLAIAISLTINEIAR 123
DB 61 GPNKKNVQNISWOPKPIINNPLFIMDISPTDICOGLGDCWLLAIGSLTTCPKLYR 120
QY 124 VVPLNGNSFOENYVGIHFQFMQYGEWEVVDRLPTKDELLFVHSABSEFMSALKE 163
DB 121 VVRGSGFKVNIHFQFMQYGEWEVVDRLPTKDELLFVHSABSEFMSALKE 180
QY 184 AVAKINGCYALSGATTEGFEFTGIAEMWELKKRPNNLFKIQALOKSGSLGCSID 243
DB 181 AVAKLSGYALSGATTEGFEFTGIAEMWELKKRPNNLFKIQALOKSGSLGCSIE 240
QY 244 ITGAADSEALTPOKLYKGAHAYSTGAEVENSNGLOKLIIRNPMGEVMTGRMNDCPS 303
DB 241 VTSDBSELSMTDKLVKGAHAYSTGAEVENSNGLOKLIIRNPMGEVMTGRMNDCPS 300
QY 304 WNTIDPERERLIRRHEDGFEFMSFSDPLRHYSRLICNLTPDILTSPTYKKMKLTMDG 363
DB 301 WEEVADIDQOLHKTEDGFEFMSYODPLNNFTLEICNLTPDILTSPTYKKMKLTMDG 360
QY 364 NWRGSGTAGGCRNYPNTFMNPOVLIKLEEDDEDEDESG---CTPLVGLIQHRRROR 420
DB 361 SWRGSGTAGGCRNYPNTFMNPOVLIKLEEDDEDEDESG---CTPLVGLIQHRRROR 420
QY 421 KMGEDMHTTIGFYVEDELSGQTNILSKNPLTNRARRSDFTINLREVLNRFKLP 480
DB 421 QOQAOIQTIGFYVAVPEKQNTODVHLKKEFTKYQDHGFSEIFTSREVSSQLRLPG 480
QY 481 EYILVSTPEPNKDGDCIRVSEKKAQVAVD-ELEANLEEDDISDDIDGVRRLFA 539
DB 481 EYILVSTPEPNKDGDCIRVSEKKAQVAVD-ELEANLEEDDISDDIDGVRRLFA 539
QY 540 QLAGDEAISAPELQTLIRVLAQRKODIKSDGSIETCKIMVMDLDDSGSGKLGKEFYI 599
DB 541 IVAEGEIGVVELQRLNLMMAIKFSFKTGRFOLDRCRCINIMDDSGSGKLGKEFYI 600
QY 600 LMTKIQKYKRIEIVDVRSGTMSYEMRKALBEAGFMPQQLHOVIVARFADQILIF 659
DB 601 LMTKIKKMDIPRECDDHSGTINSYEMRLVIEKAGIKLNKVVQVAVARADDILIF 660
QY 660 DNFRCLVLETLFKIKFOLDPENTGTIELDLISWL 695
DB 661 DSFISCFRLKLTMTFPLTMDPKNGHICLSLEQWL 696

RESULT 11
US-10-116-519-9
; Sequence 9, Application US/10116519
; Publication No. US20030114373A1
; GENERAL INFORMATION:

```
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0124 NP
; CURRENT APPLICATION NUMBER: US/10/116,519
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/281,253
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/288,768
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/296,180
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/300,620
; PRIOR FILING DATE: 2001-06-25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 709
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-116-519-9

Query Match          52.3%; Score 1943.5; DB 15; Length 709;
Best Local Similarity 51.0%; Pred. No. 9.8e-173;
Matches 354; Conservative 136; Mismatches 197; Indels 7; Gaps 4;
```

QY 7 KLAKDREAAGLSHERAIVKYNODYEALRNECLAGTLFODSPFAIPSAIGFKEIPY 66
12 RVREBDRNRGSEGTGQPLFEGQDQVYVLRCLAQKCLFEDVFPFQVVALSHSLS-Q 70
QY 67 SSKTRGWRKRPETICADPOFIIGATRTDICOGLGDCWLLAIAISLTINEIARVP 126
DB 71 KAKMKALITWKPKREICENPFIIGANRTDICOGLGDCWLLAIAISLTINEIARVP 130
QY 127 LNSFOENYVGIHFQFMQYGEWEVVDRLPTKDELLFVHSABSEFMSALKEKAYA 166
DB 131 HDOSFTEENYAGIHFQFMQYGEWEVVDRLPTKDELLFVHSABSEFMSALKEKAYA 160
QY 187 KINGCYEALSGATTEGFEFTGIAEMWELKKRPNNLFKIQALOKSGSLGCSIDITS 246
DB 191 KMGEDMHTTIGFYVEDELSGQTNILSKNPLTNRARRSDFTINLREVLNRFKLP 480
QY 247 AADSEAITFOKLYKGAHAYSTGAEVENSNGLOKLIIRNPMGEVMTGRMNDCPSMT 306
DB 251 PVGYETRMAGLYKGAHAYSTGAEVENSNGLOKLIIRNPMGEVMTGRMNDCPSMT 310
QY 307 IDPERERLIRRHEDGFEFMSFSDPLRHYSRLICNLTPDILTSPTYKKMKLTMDG 365
DB 311 VDKDEKARLQHVTEDEGFEFMSYDQVYHFTKLEICNLTPDILTSPTYKKMKLTMDG 370
QY 366 RRGSGTAGGCRNYPNTFMNPOVLIKLEEDDEDEDESGCTPLVGLIQHRRROR 425
DB 371 VRCGSGTAGGCRNYPNTFMNPOVLIKLEEDDEDEDESGCTPLVGLIQHRRROR 420
QY 426 MHTIGFYVEPELSGQTNILSKNPLTNRARRSDFTINLREVLNRFKLPGEYIIV 485
DB 431 LFTIGFYVEPELSGQTNILSKNPLTNRARRSDFTINLREVLNRFKLPGEYIIV 489
QY 486 PSTPEPNKDGDCIRVSEKKAQVAVD-ELEANLEEDDISDDIDGVRRLFA 539
DB 490 PSTPEPNKDGDCIRVSEKKAQVAVD-ELEANLEEDDISDDIDGVRRLFA 539
QY 542 QLAGDEAISAPELQTLIRVLAQRKODIKSDGSIETCKIMVMDLDDSGSGKLGKEFYI 599
DB 550 QLAGDEAISAPELQTLIRVLAQRKODIKSDGSIETCKIMVMDLDDSGSGKLGKEFYI 599
QY 602 TKIQKYKRIEIVDVRSGTMSYEMRKALBEAGFMPQQLHOVIVARFADQILIF 659
DB 610 KTIKAWQKIFKHVDTHSGTINSYEMRKALBEAGFMPQQLHOVIVARFADQILIF 659
QY 662 FVRCLVLETLFKIKFOLDPENTGTIELDLISWL 695
DB 670 FICCFVRLKLTMTFPLTMDPKNGHICLSLEQWL 703

QY 507 ADVADDELEANE-----FDISEDDIDGVRLLPQALAGDAEISAFELQTLIRVL 561
 Db 491 AITRDMGNVVIDIDPEEPKPTPPQDETEEEQFRPALFEQVAGMEVTAELVEYLVNAV 550
 QY 562 AKRODISDGSFSTICKIMVMDLSDSGKGLKEPFIYLMWKIKXOKIYREIVDVSCT 621
 Db 551 OKKODIFPKLISLISCKNITSLMOTSGVGLKEPDEFKVMKOWINLFLRFPADSGT 610
 QY 622 MNSYEMKALAEAGFKMPCOLHGYIVARFADQTLIDFNFRCLVLETLEFKIKOLDP 681
 Db 611 MSTEALATLKAFAFOJLSHLLQVILRYADELQDFDDELCLVLENAARVFOALST 670
 QY 682 ENTGTIELDL 691
 Db 671 KNEKEFIHLNI 680

RESULT 14
 US-09-768-877-24
 ; Sequence 24, Application US/09768877
 ; Patent No. US20020150896A1
 ; GENERAL INFORMATION:
 ; APPLICANT: POLONSKY, KENNETH S.
 ; APPLICANT: HORIKAWA, YUKIO
 ; APPLICANT: ODA, NAOHISA
 ; APPLICANT: COX, NANCY J.
 ; APPLICANT: SREENAN, SEAMUS
 ; APPLICANT: ZHOU, YUN-PING
 ; APPLICANT: OTANI, KENICHI
 ; APPLICANT: HANIS, CRAIG L.
 ; APPLICANT: BELL, GRAEME I.
 ; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
 ; FILE REFERENCE: ARCD:307
 ; CURRENT APPLICATION NUMBER: US/09/768,877
 ; PRIOR FILING DATE: 2001-01-23
 ; PRIOR APPLICATION NUMBER: 09/422,869
 ; PRIOR FILING DATE: 1999-10-21
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 24
 ; LENGTH: 821
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-768-877-24

Query Match 50.2%; Score 1862.5; DB 10; Length 821;
 Best Local Similarity 45.5%; Pred. No. 4.7e-165;
 Matches 347; Conservative 139; Mismatches 171; Indels 105; Gaps 8;

QY 32 YEALRNECLEAGTLFODPSFPALPSALGFKELGPSKTRGMKRPTEICADPOFTIGG 91
 Db 61 FEOHKKCLEKVLVYDPEPPDETSLFYSGKPIQ-----FWMKRPTEICENPRFTIGG 115
 QY 92 ATRFDICOGALGDCWLLAIAISLTLEBEILARVPLNOSQENYAGIFHFQWQYGEWVE 151
 Db 116 ANRTDICOGBELGDCWFLAIAIACTLNQHLFRVIPHQOSFTEYVAGIFHFQWRYGEMVD 175
 QY 152 VVVDRLPTKDGELLPFHSAGSEFMSALKEKAYAKINGCEALSGATTEGEGEDFTGCI 211
 Db 176 VVIDDCLPYNNQVLFYKSHRNEFWMSALEKAYAKHGSYALKNGATTEAMEDFTGCV 235
 QY 212 AEWEYLKPPPNLFKIIQALQKSLGCSID-----ITSADSEAITFOK----- 257
 Db 236 AEPEIHDAPSDMKIKKAIERGSLSMGCSIDCGTNNYTGSPSGLAMGELIAMVNMMD 295
 QY 258 -----LVKGHAYVTGAEBVESNGSLQKILR 283
 Db 296 NSLLQSDLDPRGSDERPTRTIIPVOYETRMACGLVRGHAYSVTGLDEVFPKGEKVLVR 355
 QY 342 IRNPMGEVEMTGRANDNOCPSMNTIDPERERLTR-HEDEGFNMSFSDPLRHSRLICN 342
 Db 356 LRNMGOVENNGMSDWMKMSFVQDKBKARLQHQVTEDEGEFMSYEDFIYHFKLEICN 415

QY 343 LTPDPTLSDTYKKMKLTKRDGNRRSGTAGCCNNYPTFMNPOYLKLEEDDEBERGE 402
 Db 416 LTRADALQSDLTQWTWVSVNGRWVRCGSCANNFPOTFWTNPOYRLKLEEDDDPDSE 475
 QY 403 SCGTFVLGLQKRRRRKRGEDMHTGPGIYVPELSQTMHLSKPNFLNRRARS 462
 Db 476 VIOSFLVLAQKRRKDRKLGASLFTTGAIYEPKEMHGNKO-HQKDFFLVNASARS 534
 QY 463 DTFINLEEVNRPFLPGEIYLVSPPEPKODGFCIRVSEKKADQOAVDEIEANTL-- 520
 Db 535 KTYINREVSQRRLRPPSEVYIVPSTIEPQOGBEFLIRVSEKRNLSSEVENTISVDRPV 594
 QY 521 -----BEFDISEDDIDG----- 533
 Db 595 KKKTKRPIIFVSDRANSKELGVQDS-EEGKGKTPSPDKQSPQPPQSSDQSEBEOQ 653
 QY 534 VRRLFAQAGEDAEISAFELQTLIRRLARQDIDSDEISFCKIMVMDLSDSGKLG 593
 Db 654 FRNIFKQIAGDDMEICADELKKVLTWNHGXKDLKTHGFTLESCKRSMIALMDTDSGKLN 713
 QY 594 LKEFYIIMTKIOKKYKOKIYREIVDRSGTMSYEMRKALBEAGFMPCOLHGYIVARFAD 653
 Db 714 LOEFHLMNKIKAMQKIFKHYVDQSGTINSYEMRANVADAGFHLNQLYDITMRADK 773
 QY 654 QLIIDFQNVFCLVRLTLEFKIKOLDPENTGTIELDLISWL 695
 Db 774 HNMIDFDSFICFVRLGEMFRFAHFDKDGDIILNVLVEML 815

RESULT 15
 US-10-116-519-7
 ; Sequence 7, Application US/10116519
 ; Publication No. US20030114373A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
 ; FILE REFERENCE: SUPERFAMILY, CAN-12 AND VARIANTS THEREOF
 ; CURRENT APPLICATION NUMBER: US/10/116,519
 ; PRIOR FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: US 60/281,253
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: US 60/288,768
 ; PRIOR FILING DATE: 2001-05-04
 ; PRIOR APPLICATION NUMBER: US 60/296,180
 ; PRIOR FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: US 60/300,620
 ; PRIOR FILING DATE: 2001-06-25
 ; NUMBER OF SEQ ID NOS: 145
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 7
 ; LENGTH: 821
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-116-519-7

Query Match 50.2%; Score 1862.5; DB 15; Length 821;
 Best Local Similarity 45.5%; Pred. No. 4.7e-165;
 Matches 347; Conservative 139; Mismatches 171; Indels 105; Gaps 8;

QY 32 YEALRNECLEAGTLFODPSFPALPSALGFKELGPSKTRGMKRPTEICADPOFTIGG 91
 Db 61 FEOHKKCLEKVLVYDPEPPDETSLFYSGKPIQ-----FWMKRPTEICENPRFTIGG 115
 QY 92 ATRFDICOGALGDCWLLAIAISLTLEBEILARVPLNOSQENYAGIFHFQWQYGEWVE 151
 Db 116 ANRTDICOGBELGDCWFLAIAIACTLNQHLFRVIPHQOSFTEYVAGIFHFQWRYGEMVD 175
 QY 152 VVVDRLPTKDGELLPFHSAGSEFMSALKEKAYAKINGCEALSGATTEGEGEDFTGCI 211
 Db 176 VVIDDCLPYNNQVLFYKSHRNEFWMSALEKAYAKHGSYALKNGATTEAMEDFTGCV 235
 QY 212 AEWEYLKPPPNLFKIIQALQKSLGCSID-----ITSADSEAITFOK----- 257

```
Db 236 AEFPEIRDPADMYKIMKKAIERGSLMGCSIDGTNMTYGTSPSGLNMGELIARVMRMD 295
QY 258 -----LVKGHAYSVTGAEEVESNGSLQKLR 283
Db 296 NSLIQSDLDLRGSDERPTRTIIPQVETRNACGLVRGHAYSVTGLDEVPRKGEKVLVR 355
QY 284 IRNPMGEVEMTGRWNDCPSWNITIDPEERERLTRR-HEDGEFMSFSDFLRHYSRLIEN 342
Db 356 LRNPMGQVEMNGSMDRWDMSFVDKDEKARLQHQVTEDEGFMSYEDFIYHFTKLEICN 415
QY 343 LTPDTLSDTYKWKLTMDGNMRGSLTAGCCRNYPNTFMNPOYLLKLEBEDDEDEGE 402
Db 416 LTAADALQSDKLQWTWVSVNEGRWVRGCSAGCRNFPDTEFTWNPQYRLKLEEDDDPDDSE 475
QY 403 SGCETFLVGLIOKRRORRORRORRORRORRORRORRORRORRORRORRORRORRORR 462
Db 476 VICSFLVALMOKNRKDKRLGASLFTIGFALYEVPEKEMHGNKQ-HLQKDFLYNASKARS 534
QY 463 DTFINLREVLNRFKLPGEYILVSTFEPNKGDFCIRVFSEKKADYQAVDEIEANL-- 520
Db 535 KTYINMRVSGRFRLPSEYIVPSTYEPHQEGEFLRVFSEKNTLSEVENTISVDRPV 594
QY 521 -----EFPDISEDIDDG----- 533
Db 595 KKKTKPIIFVSDRANSNKEIGVDQES-EEGKGKTSPPDKQKQSPQPPGSSDQSEEQQ 653
QY 534 VRRLFAOLAGDEDAISAFELQTLIRVLAKRQDIKSDGSIETCKIMVDM.LSDSGSKLG 593
Db 654 FRNIFKQIAGDDMEICADELKKVLTNVNKHDLKTHGFTLESCSRMTALMDTDGSGKLN 713
QY 594 LKEFYILMTKLOXYOKIYREIDVDRSGTMSYEMRKALEAGFKMPCQLHQVIVARFAD 653
Db 714 LQEFHLMNKKIKAMQKIFKHYDTQSGTINSYEMRNNAVNDAGFHLNNQLYDIITRYADK 773
QY 654 QLIIDFNFRCLVRLFTLFIKQIDPENTGTIELDISWL 695
Db 774 HMNIDFDSFICCFVRLGEMFRFAFDKDGDIITLNVLEWL 815
```

Search completed: July 24, 2003, 13:14:56
Job time : 50.4084 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:48:39 ; Search time 30.9606 Seconds
(without alignments)
956.622 Million cell updates/sec

Title: US-09-884-319A-7

Perfect score: 3713
Sequence: 1 MAGIAAKAKDREAEGLS.....PENTGTIELDLISMLCPSTVL 700

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3713	100.0	700	US-08-726-525-7	Sequence 7, Appli
2	3713	100.0	700	US-08-487-942-7	Sequence 7, Appli
3	3713	100.0	700	US-08-726-036A-7	Sequence 7, Appli
4	3713	100.0	700	US-09-422-869-23	Sequence 23, Appli
5	3713	100.0	700	US-09-083-516-7	Sequence 7, Appli
6	3491	94.0	700	US-09-308-345A-46	Sequence 46, Appli
7	2458.5	66.2	714	US-09-422-869-22	Sequence 22, Appli
8	2447.5	65.9	714	US-09-308-345A-47	Sequence 47, Appli
9	2370.5	63.8	703	US-08-835-099A-1	Sequence 1, Appli
10	2370.5	63.8	703	US-09-157-349-1	Sequence 1, Appli
11	2370.5	63.8	712	US-08-835-099A-2	Sequence 2, Appli
12	2370.5	63.8	712	US-09-157-349-2	Sequence 2, Appli
13	2356.5	63.5	703	US-09-422-869-27	Sequence 27, Appli
14	2308.5	62.2	703	US-09-308-345A-49	Sequence 49, Appli
15	2020	54.4	739	US-09-653-839-2	Sequence 2, Appli
16	2014	54.2	702	US-09-653-839-6	Sequence 6, Appli
17	1967	53.0	723	US-09-653-839-4	Sequence 4, Appli
18	1961	52.8	686	US-09-653-839-8	Sequence 8, Appli
19	1925.5	51.9	757	US-09-653-839-1	Sequence 1, Appli
20	1882	50.7	690	US-09-422-869-28	Sequence 28, Appli
21	1869.5	50.4	821	US-09-622-869-15	Sequence 15, Appli
22	1862.5	50.2	821	US-09-422-869-24	Sequence 24, Appli
23	1839.5	49.5	821	US-09-308-345A-48	Sequence 48, Appli
24	1594.5	34.9	828	US-09-308-345A-19	Sequence 19, Appli
25	1287	34.7	251	US-08-726-525-6	Sequence 6, Appli
26	1287	34.7	251	US-08-487-942-6	Sequence 6, Appli
27	1287	34.7	251	US-08-726-036A-6	Sequence 6, Appli

28	1287	34.7	251	4	US-09-083-516-6	Sequence 6, Appli
29	965.5	26.0	630	4	US-09-308-345A-2	Sequence 2, Appli
30	961.5	25.9	513	4	US-09-308-345A-4	Sequence 4, Appli
31	917	24.7	640	4	US-09-308-345A-7	Sequence 7, Appli
32	901.5	24.3	639	3	US-09-422-869-25	Sequence 25, Appli
33	789	21.2	648	4	US-09-308-345A-20	Sequence 20, Appli
34	780	21.0	641	3	US-09-422-869-18	Sequence 18, Appli
35	621.5	16.7	666	3	US-09-422-869-16	Sequence 16, Appli
36	607	16.3	544	3	US-09-422-869-4	Sequence 4, Appli
37	604	16.3	672	3	US-09-422-869-2	Sequence 2, Appli
38	599	16.1	517	3	US-09-422-869-8	Sequence 8, Appli
39	575	15.5	517	3	US-09-422-869-6	Sequence 6, Appli
40	531	14.3	444	3	US-09-422-869-10	Sequence 10, Appli
41	458	12.3	266	3	US-09-032-523-7	Sequence 9, Appli
42	450	12.1	268	2	US-08-835-099A-9	Sequence 9, Appli
43	450	12.1	268	3	US-09-157-349-9	Sequence 9, Appli
44	423	11.4	248	3	US-09-032-523-1	Sequence 1, Appli
45	304.5	8.2	854	2	US-08-928-692-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-08-726-525-7
; Sequence 7, Application US/08726525
; Patent No. 5789181
; GENERAL INFORMATION:
; APPLICANT: Lin, Lin-Ling
; APPLICANT: Graham, James
; TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
; TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
; TITLE OF INVENTION: BINDING
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,525
; FILING DATE: 07-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/487,942
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELEPHONE: (617) 498-8224
; TELECOMMUNICATION INFORMATION:
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 700 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-08-726-525-7
Query Match 100.0%; Score 3713; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MAGIAAKLAKDREAABGLSHERAIKYINDYBALNNECLAGTLFQDPSPFPAIPSLG 60
DB      1 MAGIAAKLAKDREAABGLSHERAIKYINDYBALNNECLAGTLFQDPSPFPAIPSLG 60
QY      61 KELGPYSSKTRGMRWRKPTETICADPOFIIGATRTDICOALGDCWLLAAIASLTNBEI 120
DB      61 KELGPYSSKTRGMRWRKPTETICADPOFIIGATRTDICOALGDCWLLAAIASLTNBEI 120
QY      121 LARVPLNOSFOENVAGIFHFQFMOYGEWVAVDDRLPTDGBELLFVHSAESESFMSAL 180
DB      121 LARVPLNOSFOENVAGIFHFQFMOYGEWVAVDDRLPTDGBELLFVHSAESESFMSAL 180
QY      181 LEKAVAKINGCYEALSGATTEGFEDEFTGIAEWYELKXPPNLFKIIQKALQKSLG 240
DB      181 LEKAVAKINGCYEALSGATTEGFEDEFTGIAEWYELKXPPNLFKIIQKALQKSLG 240
QY      241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRW 300
DB      241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRW 300
QY      301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDILTSDTYKKMKLTK 360
DB      301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDILTSDTYKKMKLTK 360
QY      361 MDGWMRSGTAGGCRNYPNTFMNPOVLIKLEBEDEDEDESGCTFLVGLIQKHRROR 420
DB      361 MDGWMRSGTAGGCRNYPNTFMNPOVLIKLEBEDEDEDESGCTFLVGLIQKHRROR 420
QY      421 KMGEDMHTIGFYIEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
DB      421 KMGEDMHTIGFYIEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
QY      481 EYILVSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEEDISEDDIDGVRRLFAQ 540
DB      481 EYILVSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEEDISEDDIDGVRRLFAQ 540
QY      541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
DB      541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
QY      601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVYVARFADQLIID 660
DB      601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVYVARFADQLIID 660
QY      661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCSVL 700
DB      661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCSVL 700

```

```

RESULT 2
US-08-487-942-7
Sequence 7, Application US/08487942
Patent No. 5817476
GENERAL INFORMATION:
APPLICANT: Lin, Lih-Ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,942
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-487-942-7

```

```

Query Match      100.0%; Score 3713; DB 2; Length 700;
Best Local Similarity 100.0%; Pred. No.0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MAGIAAKLAKDREAABGLSHERAIKYINDYBALNNECLAGTLFQDPSPFPAIPSLG 60
DB      1 MAGIAAKLAKDREAABGLSHERAIKYINDYBALNNECLAGTLFQDPSPFPAIPSLG 60
QY      61 KELGPYSSKTRGMRWRKPTETICADPOFIIGATRTDICOALGDCWLLAAIASLTNBEI 120
DB      61 KELGPYSSKTRGMRWRKPTETICADPOFIIGATRTDICOALGDCWLLAAIASLTNBEI 120
QY      121 LARVPLNOSFOENVAGIFHFQFMOYGEWVAVDDRLPTDGBELLFVHSAESESFMSAL 180
DB      121 LARVPLNOSFOENVAGIFHFQFMOYGEWVAVDDRLPTDGBELLFVHSAESESFMSAL 180
QY      181 LEKAVAKINGCYEALSGATTEGFEDEFTGIAEWYELKXPPNLFKIIQKALQKSLG 240
DB      181 LEKAVAKINGCYEALSGATTEGFEDEFTGIAEWYELKXPPNLFKIIQKALQKSLG 240
QY      241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRW 300
DB      241 SIDITSAADSEAITFOKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRW 300
QY      301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDILTSDTYKKMKLTK 360
DB      301 CPSNNTIDPEERELTRRHEDGEFMSFSDFLRHSRLICNLTPDILTSDTYKKMKLTK 360
QY      361 MDGWMRSGTAGGCRNYPNTFMNPOVLIKLEBEDEDEDESGCTFLVGLIQKHRROR 420
DB      361 MDGWMRSGTAGGCRNYPNTFMNPOVLIKLEBEDEDEDESGCTFLVGLIQKHRROR 420
QY      421 KMGEDMHTIGFYIEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
DB      421 KMGEDMHTIGFYIEVPEELSGQTNHLSKNPFLTNRARSDTFINLRVLANRFXLP 480
QY      481 EYILVSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEEDISEDDIDGVRRLFAQ 540
DB      481 EYILVSTFEENKOGDFCIRVSEKKADYQAVDEIEANLEEDISEDDIDGVRRLFAQ 540
QY      541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
DB      541 LAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGKLGKEFYIL 600
QY      601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVYVARFADQLIID 660
DB      601 WTKIQOKYKIREIDVDRSGTMNSYEMRKALBEAGFKMPCOLHQVYVARFADQLIID 660
QY      661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCSVL 700
DB      661 NFVRCIVRLFTLFIKIQOLDPENTGTIELDLISMLCSVL 700

```

RESULT 3
US-08-726-036A-7
Sequence 7, Application US/08726036A
Patent No. 5981482
GENERAL INFORMATION:
APPLICANT: Lin, Lin-Ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
TITLE OF INVENTION: BINDING
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
City: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726.036A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-08-726-036A-7

Query Match 100.0%; Score 3713; DB 2; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLADREAEAGLSHERAIKYLNDYALNECLAEAGTLFODPSFPAIPSLGF 60
1 MAGIAKLADREAEAGLSHERAIKYLNDYALNECLAEAGTLFODPSFPAIPSLGF 60
DB 1 MAGIAKLADREAEAGLSHERAIKYLNDYALNECLAEAGTLFODPSFPAIPSLGF 60
QY 61 KEIGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
61 KEIGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
DB 61 KEIGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
QY 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVVDRLPTKDGELLFVSHAEGSEFWSAL 180
121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVVDRLPTKDGELLFVSHAEGSEFWSAL 180
DB 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVVDRLPTKDGELLFVSHAEGSEFWSAL 180
QY 181 LERAVAKINCCEYALSGGATTEGFEDFTGGIAEWYELKKPPNLFKIIQALOKGSLGC 240
181 LERAVAKINCCEYALSGGATTEGFEDFTGGIAEWYELKKPPNLFKIIQALOKGSLGC 240
DB 181 LERAVAKINCCEYALSGGATTEGFEDFTGGIAEWYELKKPPNLFKIIQALOKGSLGC 240
QY 241 SIDTSAADSEAITFOQLVKGHAYSVTGABEVSNSGLQKILIRINPWGEVETGRWNN 300
241 SIDTSAADSEAITFOQLVKGHAYSVTGABEVSNSGLQKILIRINPWGEVETGRWNN 300
DB 241 SIDTSAADSEAITFOQLVKGHAYSVTGABEVSNSGLQKILIRINPWGEVETGRWNN 300
QY 301 CPSMNTIDPERERLTRRHDGEFWMSPDLRHYSLEICNLTPTDILTSDTYKKWLTG 360
301 CPSMNTIDPERERLTRRHDGEFWMSPDLRHYSLEICNLTPTDILTSDTYKKWLTG 360
DB 301 CPSMNTIDPERERLTRRHDGEFWMSPDLRHYSLEICNLTPTDILTSDTYKKWLTG 360

QY 361 MDGNWRGSTAGCCRNVPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRROR 420
361 MDGNWRGSTAGCCRNVPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRROR 420
DB 361 MDGNWRGSTAGCCRNVPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRROR 420
QY 421 KMGEDMTTGFGIYVDEEISGOTNHLNSNFFLTMRARSTFTNLREVLNRFKLP 480
421 KMGEDMTTGFGIYVDEEISGOTNHLNSNFFLTMRARSTFTNLREVLNRFKLP 480
DB 421 KMGEDMTTGFGIYVDEEISGOTNHLNSNFFLTMRARSTFTNLREVLNRFKLP 480
QY 481 EYILVSTFEPNKDGFICIVFSEKADYQAVDDEIENLEFDDISEDDIDGVRRLFAQ 540
481 EYILVSTFEPNKDGFICIVFSEKADYQAVDDEIENLEFDDISEDDIDGVRRLFAQ 540
DB 481 EYILVSTFEPNKDGFICIVFSEKADYQAVDDEIENLEFDDISEDDIDGVRRLFAQ 540
QY 541 LAGEDAISAFAEIQTLIRVLAARODIKSDGSIETCKIWNMDLSDSGSLGKKEFYIL 600
541 LAGEDAISAFAEIQTLIRVLAARODIKSDGSIETCKIWNMDLSDSGSLGKKEFYIL 600
DB 541 LAGEDAISAFAEIQTLIRVLAARODIKSDGSIETCKIWNMDLSDSGSLGKKEFYIL 600
QY 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALBEAGFMPQLHVIYARADQOLIIFD 660
601 WTKIQKQKTYREIDVDRSGTMSYEMRKALBEAGFMPQLHVIYARADQOLIIFD 660
DB 601 WTKIQKQKTYREIDVDRSGTMSYEMRKALBEAGFMPQLHVIYARADQOLIIFD 660
QY 661 NFVRCVRLTTLFKIFKQDLPENTGTIEDLISMLCFSYL 700
661 NFVRCVRLTTLFKIFKQDLPENTGTIEDLISMLCFSYL 700
DB 661 NFVRCVRLTTLFKIFKQDLPENTGTIEDLISMLCFSYL 700

RESULT 4
US-09-422-869-23
Sequence 23, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO
APPLICANT: ODA, NAOHISA
APPLICANT: COX, NANCY J.
APPLICANT: SREENAN, SEANUS
APPLICANT: ZHOU, YUN-PING
APPLICANT: OTANI, KENICHI
APPLICANT: HANIS, CRAIG I.
APPLICANT: BELL, GRAEME I.
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
FILE REFERENCE: ARCD-307
CURRENT APPLICATION NUMBER: US/09/422,869
CURRENT FILING DATE: 1999-10-21
EARLIER APPLICATION NUMBER: 60/134,175
EARLIER FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 23
LENGTH: 700
TYPE: PRT
ORGANISM: Human
US-09-422-869-23

Query Match 100.0%; Score 3713; DB 3; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLADREAEAGLSHERAIKYLNDYALNECLAEAGTLFODPSFPAIPSLGF 60
1 MAGIAKLADREAEAGLSHERAIKYLNDYALNECLAEAGTLFODPSFPAIPSLGF 60
DB 1 MAGIAKLADREAEAGLSHERAIKYLNDYALNECLAEAGTLFODPSFPAIPSLGF 60
QY 61 KEIGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
61 KEIGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
DB 61 KEIGPYSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIAISITLNEE 120
QY 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVVDRLPTKDGELLFVSHAEGSEFWSAL 180
121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVVDRLPTKDGELLFVSHAEGSEFWSAL 180
DB 121 LARVPLNOSFOENVAGIFHFQFQWYGEWVVVDRLPTKDGELLFVSHAEGSEFWSAL 180
QY 181 LERAVAKINCCEYALSGGATTEGFEDFTGGIAEWYELKKPPNLFKIIQALOKGSLGC 240
181 LERAVAKINCCEYALSGGATTEGFEDFTGGIAEWYELKKPPNLFKIIQALOKGSLGC 240
DB 181 LERAVAKINCCEYALSGGATTEGFEDFTGGIAEWYELKKPPNLFKIIQALOKGSLGC 240

QY 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
DB 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
QY 301 CPSWNTIDPEERERLIRRHEDGEFWMGSDPLRIHYSRLIECNLPDITLSTTYKKWLTJK 360
DB 301 CPSWNTIDPEERERLIRRHEDGEFWMGSDPLRIHYSRLIECNLPDITLSTTYKKWLTJK 360
QY 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
DB 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
QY 421 KMGDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPFG 480
DB 421 KMGDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPFG 480
QY 481 EYILVPSSTPEPNKQDGCIRVFSEKADYQAVNDEIEANLEEDIDSGVRLFAQ 540
DB 481 EYILVPSSTPEPNKQDGCIRVFSEKADYQAVNDEIEANLEEDIDSGVRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKLGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKLGLKEFYIL 600
QY 601 WTKIQYQKIYREIDVRSSTGMSYEMRKALEAGFKAPCOLHOVYVARFADQLIIDFD 660
DB 601 WTKIQYQKIYREIDVRSSTGMSYEMRKALEAGFKAPCOLHOVYVARFADQLIIDFD 660
QY 661 NFVRCVLRLETLFKIFKQDLPENTGTELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQDLPENTGTELDLISWLCFSVL 700

RESULT 5

US-09-083-516-7
Sequence 7, Application US/09083516
Patent No. 6300086
GENERAL INFORMATION:
APPLICANT: Lin, Lih-Ling
APPLICANT: Graham, James
TITLE OF INVENTION: NOVEL INTERLEUKIN-1 RECEPTOR
TITLE OF INVENTION: INTRACELLULAR LIGAND PROTEINS AND INHIBITORS OF LIGAND
NUMBER OF INVENTIONS: 7
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS, GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,516
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,942
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15258
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 700 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-09-083-516-7

Query Match 100.0%; Score 3713; DB 4; Length 700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGIAKLANDREAPAEGLSHERAIKYLNDYALNNECEAGTLFQDSFPALIPALGF 60
DB 1 MAGIAKLANDREAPAEGLSHERAIKYLNDYALNNECEAGTLFQDSFPALIPALGF 60
QY 61 KELGYPSSKTRGMRWRPTEICADPOITIGGARTDICOGLADQWLLAIALTLNEE1 120
DB 61 KELGYPSSKTRGMRWRPTEICADPOITIGGARTDICOGLADQWLLAIALTLNEE1 120
QY 121 LARVPLNGSFQENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSAGSEFMSAL 180
DB 121 LARVPLNGSFQENYAGIFHFQFQYGEWEVVVDRLPTKDELLFVHSAGSEFMSAL 180
QY 181 LERAYAKINGCYALSGGATTEGPEFTGIAEWEIKRPPNLFKIQALOKSILGC 240
DB 181 LERAYAKINGCYALSGGATTEGPEFTGIAEWEIKRPPNLFKIQALOKSILGC 240
QY 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
DB 241 SIDTSAADSEALTFQKLVKGHAYSVTGAEEVESNGSLQKILIRINPMGEVEMTGRWNND 300
QY 301 CPSWNTIDPEERERLIRRHEDGEFWMGSDPLRIHYSRLIECNLPDITLSTTYKKWLTJK 360
DB 301 CPSWNTIDPEERERLIRRHEDGEFWMGSDPLRIHYSRLIECNLPDITLSTTYKKWLTJK 360
QY 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
DB 361 MDGNWRGSGTAGGCRNYPNTFMNPOYLKLEEDDEDEDESGCTFLVGLIQHRRROR 420
QY 421 KMGDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPFG 480
DB 421 KMGDMHTIGFGIYEVEEELSGQTNHLSKNFPLTNRARSDFINLRVLRNFKLPFG 480
QY 481 EYILVPSSTPEPNKQDGCIRVFSEKADYQAVNDEIEANLEEDIDSGVRLFAQ 540
DB 481 EYILVPSSTPEPNKQDGCIRVFSEKADYQAVNDEIEANLEEDIDSGVRLFAQ 540
QY 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKLGLKEFYIL 600
DB 541 LAGDAEISAFELQTLIRVLAARODIKSDGFSIETCKIWMMLDSGSKLGLKEFYIL 600
QY 601 WTKIQYQKIYREIDVRSSTGMSYEMRKALEAGFKAPCOLHOVYVARFADQLIIDFD 660
DB 601 WTKIQYQKIYREIDVRSSTGMSYEMRKALEAGFKAPCOLHOVYVARFADQLIIDFD 660
QY 661 NFVRCVLRLETLFKIFKQDLPENTGTELDLISWLCFSVL 700
DB 661 NFVRCVLRLETLFKIFKQDLPENTGTELDLISWLCFSVL 700

RESULT 6

US-09-308-345A-46
Sequence 46, Application US/09308345A
Patent No. 6569665
GENERAL INFORMATION:
APPLICANT: BOEHM, Thomas;
APPLICANT: DEAR, Neil T.
TITLE OF INVENTION: No. 6569665el calpains, their preparation and use
FILE REFERENCE: 0050/47576
CURRENT APPLICATION NUMBER: US/09/308,345A
FILING DATE: 1999-05-19
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Wordperfect v. 6.1

SEQ ID NO 46
 LENGTH: 700
 TYPE: PRT
 ORGANISM: mouse
 US-09-308-345A-46

Query Match 94.0%; Score 3491; DB 4; Length 700;
 Best Local Similarity 92.9%; Pred. No. 0;
 Matches 650; Conservative 25; Mismatches 25; Indels 0; Gaps 0;

```

QY 1 MAGIAKLAKDREAAGSLGSHERAIKYLNODYEALRNECEAGTLFODPSFPAISALGF 60
DB 1 MAGIAKLAKDREAAGSLGSHERAIKYLNODYEALRNECEAGTLFODPSFPAISALGF 60
QY 61 KELGPYSSKTRGMKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASTLTNEI 120
DB 61 KELGPYSSKTRGMKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASTLTNEI 120
QY 121 LAHVPLNOSFOENYAGIFHFQWQGEWVEVVVDRLPTKDGELLFVHSAEGSEFWAL 180
DB 121 LAHVPLNOSFOENYAGIFHFQWQGEWVEVVVDRLPTKDGELLFVHSAEGSEFWAL 180
QY 181 LEKAVKINGCYALSGATTEGPEDEFTGGIAEWELKKRPNNLFKIIQALOKSLGCG 240
DB 181 LEKAVKINGCYALSGATTEGPEDEFTGGIAEWELKKRPNNLFKIIQALOKSLGCG 240
QY 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPWEVETGRANDN 300
DB 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPWEVETGRANDN 300
QY 301 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 360
DB 301 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 360
QY 361 MDNMRGSAAGCGRVNTFMNPOYLKLEEDDEDEDESGCTFVLGLOKRROR 420
DB 361 MDNMRGSAAGCGRVNTFMNPOYLKLEEDDEDEDESGCTFVLGLOKRROR 420
QY 421 KMGEDHTTIGFYIYEBEELSGQTNILSKNFTLNARERSPTFINLRVLRPKLP 480
DB 421 KMGEDHTTIGFYIYEBEELSGQTNILSKNFTLNARERSPTFINLRVLRPKLP 480
QY 481 EYLVSTPEPNKDGFCIRVFSEKKADYQAVNDEIRANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYLVSTPEPNKDGFCIRVFSEKKADYQAVNDEIRANLEEFDISDDIDGVRRLFAQ 540
QY 541 LAEDEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMDLDSGSGKLGLKEFYI 600
DB 541 LAEDEDAISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMDLDSGSGKLGLKEFYI 600
QY 601 WTKIQOKYKIREIDVDSGTNMSYEMRKALBEAGFMPCOLHQVIVARFADDLIIDF 660
DB 601 WTKIQOKYKIREIDVDSGTNMSYEMRKALBEAGFMPCOLHQVIVARFADDLIIDF 660
QY 661 NFVRCVLRLTETLFKIFKQIDPENTGTITLIDLSWLCFV 700
DB 661 NFVRCVLRLTETLFKIFKQIDPENTGTITLIDLSWLCFV 700

```

RESULT 7
 US-09-422-869-22
 Sequence 22, Application US/09422869
 Patent No. 6235481
 GENERAL INFORMATION:
 APPLICANT: POLONSKY, KENNETH S.
 APPLICANT: HORIKAWA, YUKIO
 APPLICANT: ODA, NAOHISA
 APPLICANT: COX, NAHOY J.
 APPLICANT: SREENAN, SEAMUS
 APPLICANT: ZHOU, YUN-PING
 APPLICANT: OTANI, KENICHI
 APPLICANT: HANIS, CRAIG L.
 APPLICANT: BELL, GRAEME I.

TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
 FILE REFERENCE: ARCD:307
 CURRENT APPLICATION NUMBER: US/09/422,869
 CURRENT FILING DATE: 1999-10-21
 EARLIER APPLICATION NUMBER: 60/134,175
 EARLIER FILING DATE: 1999-05-13
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 22
 LENGTH: 714
 TYPE: PRT
 ORGANISM: Human
 US-09-422-869-22

Query Match 66.2%; Score 2458.5; DB 3; Length 714;
 Best Local Similarity 63.1%; Pred. No. 1,5e-236;
 Matches 439; Conservative 124; Mismatches 130; Indels 3; Gaps 2;

```

QY 3 GIAKLAKDREAAGSLGSHERAIKYLNODYEALRNECEAGTLFODPSFPAISALGF 62
DB 13 GVAQVOKQARRELGRHENAIKYLGQVEQLRVRCLOSGTLFRDEAFPPVQSIGYKD 72
QY 63 LGPYSSKTRGMKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASTLTNEI 122
DB 73 LGPYSSKTRGMKRPTEICADPOFIIGATRTDIOGALGDCWMLAALASTLTNEI 122
QY 123 RVPPLNOSFOENYAGIFHFQWQGEWVEVVVDRLPTKDGELLFVHSAEGSEFWAL 182
DB 123 RVPPLNOSFOENYAGIFHFQWQGEWVEVVVDRLPTKDGELLFVHSAEGSEFWAL 182
QY 183 KAVAKINGCYALSGATTEGPEDEFTGGIAEWELKKRPNNLFKIIQALOKSLGCG 242
DB 183 KAVAKINGCYALSGATTEGPEDEFTGGIAEWELKKRPNNLFKIIQALOKSLGCG 242
QY 243 DITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPWEVETGRANDN 302
DB 243 DITSAADSEAITFOKLKVGHAYSVTGAEEVESNGSLQKILIRINPWEVETGRANDN 302
QY 303 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 362
DB 303 CPBWNITIDPEERELTRHEDGEFWMSPDFLRHYSRLICNLTPDTLTSPTYKKMKLT 362
QY 363 GNRGSAAGCGRVNTFMNPOYLKLEEDDEDE--DGSGCTFVLGLOKRROR 420
DB 363 GNRGSAAGCGRVNTFMNPOYLKLEEDDEDE--DGSGCTFVLGLOKRROR 420
QY 421 KMGEDHTTIGFYIYEBEELSGQTNILSKNFTLNARERSPTFINLRVLRPKLP 480
DB 421 KMGEDHTTIGFYIYEBEELSGQTNILSKNFTLNARERSPTFINLRVLRPKLP 480
QY 481 EYLVSTPEPNKDGFCIRVFSEKKADYQAVNDEIRANLEEFDISDDIDGVRRLFAQ 540
DB 481 EYLVSTPEPNKDGFCIRVFSEKKADYQAVNDEIRANLEEFDISDDIDGVRRLFAQ 540
QY 540 QLAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMDLDSGSGKLGLKEFYI 599
DB 540 QLAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMDLDSGSGKLGLKEFYI 599
QY 599 QLAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMDLDSGSGKLGLKEFYI 612
DB 599 QLAGEDAEISAFELQTLIRVLAKRQDIKSDGFSIETCKIWMDLDSGSGKLGLKEFYI 612
QY 612 LMRIRINYLSIFRKFPLDGSWSAYEMMAIESAFKINKKLYEILTRYSEPDIAVDF 672
DB 612 LMRIRINYLSIFRKFPLDGSWSAYEMMAIESAFKINKKLYEILTRYSEPDIAVDF 672
QY 672 DNFVRCVLRLTETLFKIFKQIDPENTGTITLIDLSWLCFV 708
DB 672 DNFVRCVLRLTETLFKIFKQIDPENTGTITLIDLSWLCFV 708

```

RESULT 8
 US-09-308-345A-47
 Sequence 47, Application US/09308345A
 Patent No. 656965
 GENERAL INFORMATION:
 APPLICANT: BOEHM, THOMAS;

APPLICANT: DEAR, Neil T.
TITLE OF INVENTION: No. 656965e1 calpains, their preparation and use
FILE REFERENCE: 0050/47576
CURRENT APPLICATION NUMBER: US/09/308,345A
NUMBER OF FILING DATE: 1999-05-19
NUMBER OF SEQ ID NOS: 49
SOFTWARE: WordPerfect v. 6.1
SEQ ID NO: 47
LENGTH: 714
TYPE: PRT
ORGANISM: human
US-09-308-345A-47

Query Match 65.9%; Score 2447.5; DB 4; Length 714;
Best Local Similarity 62.9%; Pred. No. 1.9e-235;
Matches 438; Conservative 124; Mismatches 131; Indels 3; Gaps 2;

3 GIATLAKDRRAAGLSGHERAIKYLNDYALNECEIAGTLQODSEFPALPSALGFE 62
13 GVSAGVQKORARELGLGRHENAIXYLGDYEQRLRVCLOSGLTLPDEAFPPVPSLGKD 72
63 LGPSSKTRGMRKRPTEICADPOFIIGATRTDICOAGLDCWLLAIAASLTINEEIA 122
73 LGPSSKTYGKMKRPTEILSNPOFIYDGAIRTDICQALDCWLLAIAASLTINDLILH 132
123 RVVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFMSALIE 182
133 RVVHGSFQNGYAGIFHFQFQYGEWVVDLPLKOKLTVHSAEENEFMSALLE 192
183 KAYKINGCYBALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILGCSI 242
193 KAYKAVNGSYBALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILGCSI 252
243 DITSAADSEATPCKLVKGAHYVTGAEEVNSGLQTLIRHPMGVEWGTGRMNCP 302
253 DISVLDEALTEFKLVKGAHYVTGAQVYRGQVSLIMANPMGEWGTGMSDSS 312
303 SWNTIDEEERRLTRRHEDGEFMSFSDFLHYSRLICNLTPDLTSDTYKKWKLTM 362
313 EMNVDPYERDQLAVKMGDEGFMSFRDPMREPTRLICNLTPDALSRTIRKNTTIE 372
363 GNMRRGSTAGCCNRYPTFMNPOYLKLEEDDEB--DESGTFLVGLIOKRRR 420
373 GPMRRGSTAGCCNRYPTFMNPOYLKLEEDDEB--DESGTFLVGLIOKRRR 432
421 KMGEDMHTIGFQIYEVPELSGQTNHLSKNFPLTRNRSRDPFINREVLNPKLPG 480
433 RFGDMETIGFAVEPELVGQPAVHLKDFFLANASRANSEOFINREVTFRLLPG 492
481 EYILVSTFEENKDGDFCIRVSEKADYQAVDEIEANL--EEFDISEDDIDGVRLLFA 539
493 EYVVPSTFEENKDGDFLRFSEKAGTVELDQIQANLPDEQLSEEDIDENPKLFR 552
540 QLAEDDAISAFELQTLIRVLAKQDIXSGFSETECKIVMDLSDSGSKLKEFYI 599
553 QLAEDDAISAFELQTLIRVLAKQDIXSGFSETECKIVMDLSDSGSKLKEFYI 612
600 LMTIOCKOKIYREIDVDRSGTMSYEMRKALKEAGFPMCOLHOVIYARPADQLIIDF 659
613 LMTIRKTYLSIFRKFPLDCKSGMSAYEMRKALKEAGFPMCOLHOVIYARPADQLIIDF 672
660 DNFRVCLVRLLETLFKIFKQDLPENTGTIELDLISML 695
673 DNFRVCLVRLLETLFKIFKQDLPENTGTIELDLISML 708

RESULT 9
US-08-835-099A-1
Sequence 1, Application US/08835099A
Patent No. 5874277
GENERAL INFORMATION:
APPLICANT: SHINTANI, Yasuehi
APPLICANT: NISHI, Kazuori

APPLICANT: KAWAMOTO, Tomohiro
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
FILE REFERENCE: AND USE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,099A
FILING DATE: 04-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 083649/1996
FILING DATE: 05-APR-1996
APPLICATION NUMBER: 97105508.2
FILING DATE: 03-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 703 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-835-099A-1

Query Match 63.8%; Score 2370.5; DB 2; Length 703;
Best Local Similarity 61.5%; Pred. No. 9.2e-228;
Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;

1 MAGIAKLADREABESLSGHERAIKYLNDYALNECEIAGTLQODSEFPALPSALGFE 60
1 MAAGVAGVSRORATQGLSNQNALKYLGDFFTLRQOCLDSVLFKDFEPACPSALGY 60
61 KELGYPSSKTRGMRKRPTEICADPOFIIGATRTDICOAGLDCWLLAIAASLTINEE 120
61 KDLGPGSPQTOGIIWKRPTEICPSPOFIYDGAIRTDICQGLDCWLLAIAASLTINEE 120
121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFMSAL 180
121 LYRVPRDQDFQENYAGIFHFQFQYGEWVVDRLPTKDELLFVHSAESSEFMSAL 180
181 LERAYAKINGCYBALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILG 240
181 LERAYAKINGCYBALSGGATTEGFEPTGIAEWYELKRPPLFKIIOKALOKSILG 240
241 SIDTSAADSEATPCKLVKGAHYVTGAEEVNSGLQTLIRHPMGVEWGTGRMNCP 300
241 SIDVSSAAERALTISQTLVSHAYVTGVEVAFQHPETLIRHPMGVEWGTGMSDSD 300
301 CPMNTIDEEERRLTRRHEDGEFMSFSDFLHYSRLICNLTPDLTSDTYKKWKLTK 360
301 APEWNHIDPRKKELDKVEDGEFMSFSDFLHYSRLICNLTPDALSRTIRKNTTIE 360
361 MDGMRGSTAGCCNRYPTFMNPOYLKLEEDDEB--DESGTFLVGLIOKRRR 418
361 FNGHWRGSTAGCCNRYPTFMNPOYLKLEEDDEB--DESGTFLVGLIOKRRR 420

Best Local Similar

```

Matches      430; Conservative      123; Mismatches      143; Indels      3; Gaps
QY      1 MAGTAAKXAKDREAAEGSGSHERRAKTYKYNODYEALRNELTACGTPODSFPAIPALAF 60
Dd      1 MAAQAAGS/RKQAAVQGLGSSNQNAKXVIGQDCKTLRQOCLDSGVLFKDEPPPCFALG 60
QY      61 KELGFSKRTGRMKWRKPTETICADPOFTIGGATRTDTCGALGDCMLLAIASTLTNEET 120
Dd      61 KDLGGSPQTQGIIMKRTELCPSPQFIVGATRTDTCGGAGDCMLLAIASTLTNEEL 120
QY      121 LARVPLNQSTQENYAGIFHFQFQWQYGEWVEYVVDRLPTKQGEILFVHSARGSEWSAL 180
Dd      121 LYRVFPRQDQENYAGIFHFQFQWQYGEWVEYVVDRLPTKQGEILFVHSARGSEWSAL 180
QY      181 LEXYAKIKNGCYEALSGGATTEGFEDPFGTIAEWTELKRPNTLFFKIOKALQKSLGC 240
Dd      181 LEKAYAKIKNGCYEALSGGSTYBGFEDPFGTISEFTDLKKPPNLTQIIRKALCAGSLGC 240
QY      241 SIDITSAADSEALITFOKLKVGHAYSVTGAEEVSENGSLQKLIRINPWEVEMWTGRMND 300
Dd      241 SIDVSAAEALITISQKLKVSHAYSVTGAEEVNFQGHPEKTLIRLNPMGEVEMSGAMSD 300
QY      301 CPSWNTIDPEERELTRHEDGEEFMSSDPLRHSRLIECNLTPTDILSDTYKKYKLT 360
Dd      301 APEWNHIDPRKEELDKVEDEGEFMSLSDFVRQSRLEICNLSPDSLSSEVHKWNLVL 360
QY      361 MDGNWRSGSTAGGCNNYENPTFMNPNQYLKLEEDDEED--GESGCTFLVGLIQGRR 418
Dd      361 FNGHMTIRSTAGGCNNYPATYTNQFKIRLDEVEDDEESIGECCTYLLGLMQNRKM 420
QY      419 QRKQGEDMHTIGFYIEVPEELSGQTNILHSKNFFLTNRARSDPTFNLREVLFKLP 478
Dd      421 RKRIGQGMLSIGYAVYQVPEKELESTDLHLGDFFLAVQPSARTSTYVNLREVSGRARLP 480
QY      479 PGEYIIVSTEEPNNQDPCRTIRVPESEKKADYQAVDDEIENLEEDDIS--DIDGCVRL 537
Dd      481 PGEYIIVSTIEPFDGFCFLRVSEKKAQALIEIDVVAAGNPYEPHPSVVDDEDDFRFL 540
QY      538 FAQIAGEDAEISAFELQTLIRVLAKRQDIDKSDGFSIETCKIMVMDLSDSGSKGLKEEF 597
Dd      541 FEKLAKGQSEITANMLKILNEAFESKRDIKPDGFINIMTCRMISILDSNGTIGLGAVEF 600
QY      598 YILMTKIOKQYKIREIDVDSGTMNSYEMRKALAEAGKMPCOLHOYVAPADQOLI 657
Dd      601 KTLMLKIOKQYKIRIWEYDYNHSGTIDAEHMRALAKAGFTJNSQVOQTALRYACSKGI 660
QY      658 DEDNFRCLVRLTTLFKIFQOLDPENTGTIETDLISWLC 696
Dd      661 NFDSPVACMIRLETLFKLFLSLDEBDKQMGVQSLAEWLC 699

RESULT 11
US-08-835-099A-2
: Sequence 2, Application us/08835099A
: Patent No. 5874277
:
: GENERAL INFORMATION:
:   APPLICANT: SHINTANI, Yasushi
:   APPLICANT: NISHI, Kazuo
:   APPLICANT: KAWAMOTO, Tomohiro
:   TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
:   TITLE OF INVENTION: AND USE
:   NUMBER OF SEQUENCES: 18
:   CORRESPONDENCE ADDRESS:
:   ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
:   STREET: 130 Water Street
:   CITY: Boston
:   STATE: MA
:   COUNTRY: USA
:   ZIP: 02109
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Diskette
:   COMPUTER: IBM Compatible
:   OPERATING SYSTEM: DOS

```



```

Db 190 LERAYALNGCYELAGSTVEGEDEFTGISEFYDKKRPANLYQIIRKLCAGSLGCG 249
Qy 241 SIDDITSAADSEALTFQKLVKGHAYSTGAEEVENSNGLOKLIIRNRNMGVEWETGRANDN 300
Db 250 SIVVSAABEAATISQULVKGHAYSTGVEEVNQSGPEKILIRNRNMGVEWESGANSDD 309
Qy 301 CPSEMTIDPEBERELTRRHEDGEFWMSPDFLRHYSLEICNLTPDITLSDTYKKWLTG 360
Db 310 APEMNHIDPRKEELDKKVEDGEFWMSPDFVRQFSRLICNLSPDLSSEBVKHKNLV 369
Qy 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEDEDEBDG--ESGCTPLVGLIOKRRR 418
Db 370 FNGHWRGSTAGGCRNYPNTFMNPOYLKLEDEDEBDG--ESGCTPLVGLIOKRRR 429
Qy 419 ORKMGEDMTIGFIYVPEELSGQTNHLSKNPFLTNRARESDPTINLREVLNRPKLP 478
Db 430 KRKIGQMLSTIGYAVVQPKLESHDHALGRDFFLAYOPSARTSTYVNLREVSGRRLP 489
Qy 479 PGEYIIVPSTFEENKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 537
Db 490 PGEYIIVPSTFEENKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 549
Qy 538 PAOLAGDAEISAELOTLIRRVLAQRDIDSQFSIETCKINWMDSDSGKLGK 597
Db 550 FEKLAGDSEITANALKILNEAFSKRTDIKFDEFINTCREMISLSDSGTGLGAVEF 609
Qy 598 YIIMTKIOKYKIREIDVDRSGTMSYEMRKALBEAGFKMPCOLHOVIYARFADOL 657
Db 610 KTMKLIKQKLEIWEVDYHSGTIDHNEHTARKKAGFTLNSGOQTILARVACSKLGI 669
Qy 658 DFDNFRCLVRLTFLFKIFKQDPEBNTGIELDLSWLC 696
Db 670 NFDSPVACMIRLETFLFLDDEDKDQWQSLAEWLC 708

RESULT 13
US-09-422-869-27
; Sequence 27, Application US/09422869
; Patent No. 6235481
; GENERAL INFORMATION:
; APPLICANT: POLONSKY, KENNETH S.
; APPLICANT: HORIKAWA, YUKIO
; APPLICANT: ODA, NAOMISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BEIL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/422,869
; EARLIER APPLICATION NUMBER: 1999-10-21
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 703
; TYPE: PRT
; ORGANISM: RAT
US-09-422-869-27

```

```

Query Match 63.5%; Score 2356.5; DB 3; Length 703;
Best Local Similarity 60.9%; Pred. No. 2,3e-226;
Matches 427; Conservative 127; Mismatches 140; Indels 7; Gaps 3;
Qy 1 MAGIAAKLAKDREAAGLSHERAIXYLNODYALNECIEAGTLFODSPFPAIPSLG 60
Db 1 MALIAGVSKQRAVABGLSNOAVKYLGDFFTLRKQCLNSGVLFKDPFPACPSALG 60
Qy 61 KEIGPYSSKTRGRMKRPTEICADPOFIIGATRTDICGALDPCWLLAAIASLTINEE 120

```

```

Db 61 KDLGPGSPDIOGIVWKRPTELCPNPQFIYVGAATRTDIOGGLDPCWLLAAIASLTINEE 120
Qy 121 LARVPLNOSFOBNVYGIFFHOPOWGEWVVDRLPRKDELLFVHABEGSEFSA 180
Db 121 LRVLPDQSFQKQVAGIIFHOFQWGEWVVIDLPLPKNOQLFHSBEGSEFSA 180
Qy 181 LERAYALNGCYELAGSTVEGEDEFTGISEFYDKKRPANLYQIIRKLCAGSLGCG 240
Db 181 LERAYALNGCYELAGSTVEGEDEFTGISEFYDKKRPANLYQIIRKLCAGSLGCG 240
Qy 241 SIDDITSAADSEALTFQKLVKGHAYSTGAEEVENSNGLOKLIIRNRNMGVEWETGRANDN 300
Db 241 SIVVSAABEAATISQULVKGHAYSTGVEEVNQSGPEKILIRNRNMGVEWESGANSDD 300
Qy 301 CPSEMTIDPEBERELTRRHEDGEFWMSPDFLRHYSLEICNLTPDITLSDTYKKWLTG 360
Db 301 APEMNHIDPRKEELDKKVEDGEFWMSPDFVRQFSRLICNLSPDLSSEBVKHKNLV 360
Qy 361 MDGNWRGSTAGGCRNYPNTFMNPOYLKLEDEDEBDG--ESGCTPLVGLIOKRRR 418
Db 361 FNGHWRGSTAGGCRNYPNTFMNPOYLKLEDEDEBDG--ESGCTPLVGLIOKRRR 420
Qy 419 ORKMGEDMTIGFIYVPEELSGQTNHLSKNPFLTNRARESDPTINLREVLNRPKLP 478
Db 421 KRKIGQMLSTIGYAVVQPKLESHDHALGRDFFLAYOPSARTSTYVNLREVSGRRLP 480
Qy 479 PGEYIIVPSTFEENKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 535
Db 481 PGEYIIVPSTFEENKGDPCIRVSEKKADYQAVDEIEANLEBFDISE--DDIDDGVRL 538
Qy 536 RLPALAGDAEISAELOTLIRRVLAQRDIDSQFSIETCKINWMDSDSGKLGK 595
Db 539 SLEEFYKQSEISANQKRVINBYLSKRDIMFDFGNITCREMISLSDSGTGLGAVEF 598
Qy 596 EYIIMTKIOKYKIREIDVDRSGTMSYEMRKALBEAGFKMPCOLHOVIYARFADOL 655
Db 599 EFKTLMKIRTYLEIPEOMDHNHVGITIEAHEMTALKAGFTLNSGOQTILARVACSKL 658
Qy 656 IIDEDNFRCLVRLTFLFKIFKQDPEBNTGIELDLSWLC 696
Db 659 GVDNFRCLVRLTFLFKIFKQDPEBNTGIELDLSWLC 699

RESULT 14
US-09-308-345A-49
; Sequence 49, Application US/09308345A
; Patent No. 6569665
; GENERAL INFORMATION:
; APPLICANT: BOEHM, Thomas;
; APPLICANT: DEAR, Neil T.
; TITLE OF INVENTION: No. 6569665el calpains, their preparation and use
; FILE REFERENCE: 0050/4/576
; CURRENT APPLICATION NUMBER: US/09/308,345A
; EARLIER APPLICATION NUMBER: 1999-05-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: WordPerfect v. 6.1
; SEQ ID NO 49
; LENGTH: 703
; TYPE: PRT
; ORGANISM: rat
US-09-308-345A-49

```

```

Query Match 62.2%; Score 2308.5; DB 4; Length 703;
Best Local Similarity 60.1%; Pred. No. 1,5e-221;
Matches 421; Conservative 125; Mismatches 148; Indels 7; Gaps 3;
Qy 1 MAGIAAKLAKDREAAGLSHERAIXYLNODYALNECIEAGTLFODSPFPAIPSLG 60
Db 1 MALIAGVSKQRAVABGLSNOAVKYLGDFFTLRKQCLNSGVLFKDPFPACPSALG 60
Qy 61 KEIGPYSSKTRGRMKRPTEICADPOFIIGATRTDICGALDPCWLLAAIASLTINEE 120

```

```

Db      61 KDLGPGSPDTGIGIWMKPTLCPNPQFIVGATRTDTRGGLVDCWMLAAIASLTLEK 120
Qy      121 LARVPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAESEFMSAL 180
Db      121 LVRPLPRDQSQKQYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAESEFMSAL 180
Qy      181 LEKAYKINGCYEALSGGATTEGFEDFTGIAEWELKKPPNLFKIIQKALQKSLGCG 240
Db      181 LEKAYKINGCYEALSGGATTEGFEDFTGIAEWELKKPPNLFKIIQKALQKSLGCG 240
Qy      241 SIDTSAADSEAITFOKLYKHAIVSYTGAEEVENSGLQKIRIRNPMGEVETGRANDN 300
Db      241 SIDTSAADSEAITFOKLYKHAIVSYTGAEEVENSGLQKIRIRNPMGEVETGRANDN 300
Qy      301 CPSNWTIDPEERERLTRRHEDEGFPMSPDFLRHYSRLICNLTPDITLSDTYKKWLT 360
Db      301 CPSNWTIDPEERERLTRRHEDEGFPMSPDFLRHYSRLICNLTPDITLSDTYKKWLT 360
Qy      361 MDGNWRGSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--ESGCTFLVGLIQKHRR 418
Db      361 MDGNWRGSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--ESGCTFLVGLIQKHRR 418
Qy      419 QKRGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRAERBDTFLNLEVNRFCLP 478
Db      419 QKRGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRAERBDTFLNLEVNRFCLP 478
Qy      479 PGEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLE--FDISEDDIDGVR 535
Db      479 PGEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLE--FDISEDDIDGVR 535
Qy      536 RLFAQLAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGSKGLK 595
Db      536 RLFAQLAGEDAEISAFELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGSKGLK 595
Qy      596 EFLYLMKIRTYLEFQEMDNHVTIEAHHERKALKKAGPTLNNQVQOTIAMRYACSL 658
Db      596 EFLYLMKIRTYLEFQEMDNHVTIEAHHERKALKKAGPTLNNQVQOTIAMRYACSL 658
Qy      656 IIDFNFVRLVLETFLFKIFKQDPEWNTGTELDLSMLC 696
Db      656 IIDFNFVRLVLETFLFKIFKQDPEWNTGTELDLSMLC 696
RESULT 15
US-09-653-839-2
; Sequence 2, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
; FILE REFERENCE: LEX-0038-USA
; CURRENT APPLICATION NUMBER: US/09/653, 839
; PRIORITY FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/152, 057
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 739
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-653-839-2

```

Query Match 54.4%; Score 2020; DB 4; Length 739;
 Best Local Similarity 52.5%; Pred. No. 1,1e-192;
 Matches 366; Conservative 134; Mismatches 193; Indels 4; Gaps 2;

```

Qy      3 GIAKLAKDREAAGLGSHERAITYLNODYEALRNCELEAGTLFODSPFPAISALGFE 62
Db      37 GMAVHINNSLKKAKGVQGHNAQNFQNSFEELRAACLRGELFEDLPFAEPPSSLGFKD 96
Qy      63 LGFYSSKTRGMKRPPEITCADPOFIIGATRTDTCGALGDCWMLAAIASLTLEK 122
Db      97 LGFNSKNVQNIISWQRPDIINPLFTMDGISPFIDICGILGDCWMLAAIASLTTCPLLY 156
Qy      123 RVPPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAESEFMSAL 182
Db      157 RVPPLNOSFQENYAGIFHFQFQWYGEWVVDRLPTKDGELLFVHSAESEFMSAL 182
Qy      183 KAYAKINGCYEALSGGATTEGFEDFTGIAEWELKKPPNLFKIIQKALQKSLGCG 242
Db      217 KAYAKINGCYEALSGGATTEGFEDFTGIAEWELKKPPNLFKIIQKALQKSLGCG 242
Qy      243 DITSADSEAITFOKLYKHAIVSYTGAEEVENSGLQKIRIRNPMGEVETGRANDN 302
Db      277 EVTSDSLEBEMTKMLVRGHAIVSYTGAEEVENSGLQKIRIRNPMGEVETGRANDN 302
Qy      303 SMTIDPEERERLTRRHEDEGFPMSPDFLRHYSRLICNLTPDITLSDTYKKWLT 362
Db      337 EMBEVASDIOMQLHKTEDGFPMSPDFLRHYSRLICNLTPDITLSDTYKKWLT 362
Qy      363 GNNRSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--CTFLVGLIQKHRRQ 419
Db      397 GNNRSGTAGGCRNYPTFMNPOYLKLEEDDEDEDEG--CTFLVGLIQKHRRQ 419
Qy      420 RKGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRAERBDTFLNLEVNRFCLP 479
Db      457 RKGEDMHTTGFQIYEVPEELSGQTNHLSKNFPLNRAERBDTFLNLEVNRFCLP 479
Qy      480 GEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLE--FDISEDDIDGVR 538
Db      517 GEYILVSTPEPNKDGFCIRVFSEKKADYQAVDEIEANLE--FDISEDDIDGVR 538
Qy      539 AOLAGEDEAISAELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGSKGLKEFY 598
Db      577 AOLAGEDEAISAELQTLIRVLAKRODIKSDGFSIETCKIMVMDLSDSGSKGLKEFY 598
Qy      599 ILMTKIQKTYREIDVDSGTNNSYEMKALBEAGFKRPPCQLOHYIVARPDQOL 658
Db      637 ILMTKIQKTYREIDVDSGTNNSYEMKALBEAGFKRPPCQLOHYIVARPDQOL 658
Qy      659 FDNFVRLVLETFLFKIFKQDPEWNTGTELDLSMLC 695
Db      697 FDNFVRLVLETFLFKIFKQDPEWNTGTELDLSMLC 695

```

Search completed: July 24, 2003, 12:58:04
 Job time : 32.9606 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 12:27:29 ; Search time 89.9793 Seconds
(without alignments)
1234.824 Million cell updates/sec

Title: us-09-884-319a-7

Perfect score: 3713
Sequence: 1 MAGIAAKLANDREAREGLS.....PENTGTIELDLISWLCFSVL 700

Scoring table: BL0SUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_190jun03: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3713	100.0	700	AAW19992	Human CAMP used to
2	3713	100.0	700	AA37797	Human Interleukin-
3	3713	100.0	700	AA25059	Human calpain prot
4	3694	99.5	700	AA86128	Human calpain 80kD
5	3518	94.7	700	AA86130	Rat calpain 80kDa
6	2435	65.6	713	AA23085	Calcium-activated
7	2370.5	63.8	703	AAW1554	Human calpain. Ho
8	2370.5	63.8	712	AAW1565	Human calpain. Ho
9	2367.5	63.8	703	AAE14358	Human protease PR

10	2367.5	63.8	703	23	AAU72884	Human aspartyl pro
11	2020	54.4	739	22	AA62152	Novel human protei
12	2014	54.2	702	22	AAW79025	Human protein SEQ
13	2014	54.2	702	22	AA62154	Novel human protei
14	2014	54.2	702	22	AA64595	Human testis-speci
15	2014	54.2	702	23	AAU72887	Human aspartyl pro
16	2012	54.2	702	21	AA840701	Human ORFX ORF465
17	2002	53.9	787	22	AA812017	Human calpain homo
18	2002	53.9	787	22	AA800009	Human protein SEQ
19	1967	53.0	723	22	AA62153	Novel human protei
20	1961	52.8	666	22	AA62155	Novel human protei
21	1925.5	51.9	757	20	AA730342	A calpain protein
22	1889.5	50.9	773	24	ABU52620	Human NOVX protei
23	1888.5	50.7	767	23	AAE19164	Human protease, PR
24	1882	50.7	690	24	ABU6304	Human CAN-12 assoc
25	1862.5	50.2	821	17	AA895979	Calpain large subu
26	1610.5	43.4	791	22	AB63592	Drosophila melanog
27	1594.5	42.9	828	22	AB63528	Drosophila melanog
28	1532.5	41.3	1069	22	AB626746	Novel human diag
29	1495.5	40.3	399	22	AAW41149	Human polypeptide
30	1453	39.1	720	23	AAW49717	Murine calpain pro
31	1354	36.5	565	22	AB66731	Drosophila melanog
32	1335	36.0	708	23	AAE19179	Human protease, PR
33	1335	36.0	708	23	AAU72885	Human aspartyl pro
34	1328.5	35.8	648	23	ABG70268	Human Calpain-like
35	1287	34.7	251	18	AAW19991	Type I, p80 IL-1-r
36	1287	34.7	251	18	AAW37796	Human interleukin-
37	1218	32.8	300	22	AAW39363	Human polypeptide
38	1150	31.0	697	24	ABU6328	Human calpain cyst
39	1148.5	30.9	694	24	ABU6327	Human calpain cyst
40	1134	30.5	711	23	AAU72886	Human aspartyl pro
41	1126.5	30.3	343	22	AAE09757	Calpain cysteine p
42	1120.5	30.2	462	23	AAW49719	Murine calpain pro
43	1120.5	30.2	518	23	AAW49718	Murine calpain pro
44	1111.5	29.9	659	24	ABU6375	Human CAN-12v2 N-t
45	1110	29.9	656	24	ABU6373	Human CAN-12v1 N-t

ALIGNMENTS

RESULT 1
AAW19992 standard; Protein; 700 AA.
ID AAW19992: *
AC AAW19992: *
XX
DT 27-AUG-1997 (first entry)
XX
DE Human CAMP used to identify inhibitors of interleukin-1 activity.
XX
KW IL-1; interleukin; receptor; ligand; screening assay; inhibitor;
KW IL-1 mediated response; inflammation; inflammatory; antibody;
KW Intracellular domain; CAMP; calcium activated neutral protease.
XX
OS Homo sapiens.
XX
PN WO9640907-A1.
XX
PD 19-DEC-1996.
XX
PF 06-MAY-1996; 96MO-US06363.
XX
PR 07-JUN-1995; 95US-0487942.
XX
XX (GENY) GENETICS INST INC.
XX Graham J, Lin L;
PI WPI; 1997-052315/05.
DR
XX Interleukin-1 receptor intracellular ligand proteins and related DNA
XX - used to identify inhibitors of the proteins for treatment of
PT

PT inflammation
 XX
 PS Claim 14; Page 36-38; 54pp. English.
 XX
 CC AAM1992 represents human calcium activated neutral protease (CANP).
 CC This protein was found to have an area of high homology with an
 CC interleukin-1 receptor (IL-1-R) intracellular ligand (encoded by cDNA
 CC clone 14w, see AAT71218) and thus will display some of the same
 CC properties of this protein. IL-1-R intracellular ligand proteins are
 CC used to screen for agents (e.g. antibodies) that are capable of
 CC inhibiting or blocking the binding of an IL-1-R intracellular ligand
 CC to the intracellular domain of IL-1-R, i.e. inhibitors of IL-1
 CC activity. Such agents can be used to treat inflammatory conditions.
 XX
 SQ Sequence 700 AA;
 Query Match 100.0%; Score 3713; DB 18; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGIAAKLAKDREAEGLSHERAIKYNDYALNRECEAGTLFODSPFPAIPSLGF 60
 DB 1 MAGIAAKLAKDREAEGLSHERAIKYNDYALNRECEAGTLFODSPFPAIPSLGF 60
 QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAISLTNBEI 120
 DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAISLTNBEI 120
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTXGELLFVHSAEGSEFWAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTXGELLFVHSAEGSEFWAL 180
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTXGELLFVHSAEGSEFWAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTXGELLFVHSAEGSEFWAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKRPPLFKIIQKALQKSLGCG 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKRPPLFKIIQKALQKSLGCG 240
 QY 241 SIDITSAADSEAITFOKLVKHAIVTGAEEVNSGLQKIRIRNPGVEWMTGRWNN 300
 DB 241 SIDITSAADSEAITFOKLVKHAIVTGAEEVNSGLQKIRIRNPGVEWMTGRWNN 300
 QY 301 CPSWNTIDPERERELTRHDEGEFWMSPDLRHYSRLIECNLPDILTSTYKKMKLTK 360
 DB 301 CPSWNTIDPERERELTRHDEGEFWMSPDLRHYSRLIECNLPDILTSTYKKMKLTK 360
 QY 361 MDGMMRGSTAGGCRNYPNTFMMNPQYIKLEBEDEDEDESGCTFLVGLIOKRRROR 420
 DB 361 MDGMMRGSTAGGCRNYPNTFMMNPQYIKLEBEDEDEDESGCTFLVGLIOKRRROR 420
 QY 421 KMGEDMHTIGFGIYEPBELSGQTNILSKNPFLLNRAERSDTFINLREVLNRPKLP 480
 DB 421 KMGEDMHTIGFGIYEPBELSGQTNILSKNPFLLNRAERSDTFINLREVLNRPKLP 480
 QY 481 EYIIVPSTFEENKGGDFCIRVFSEKKADYQAVDEIEANLESPDISEDIDGVRRLFAQ 540
 DB 481 EYIIVPSTFEENKGGDFCIRVFSEKKADYQAVDEIEANLESPDISEDIDGVRRLFAQ 540
 QY 541 LAGSDAISIAPFLOTILRRVLAARODIKSDGFSIETCKIWMYMDSDSGKLGKEFYIL 600
 DB 541 LAGSDAISIAPFLOTILRRVLAARODIKSDGFSIETCKIWMYMDSDSGKLGKEFYIL 600
 QY 601 WTKIQTKYIREIDVDSGTMSYEMRKALBEAGFKMPCQLHQYVARFADQIILDFD 660
 DB 601 WTKIQTKYIREIDVDSGTMSYEMRKALBEAGFKMPCQLHQYVARFADQIILDFD 660
 QY 661 NFVACIVRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
 DB 661 NFVACIVRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
 QY 661 NFVACIVRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700
 DB 661 NFVACIVRLETLFKIFKQIDPENTGTIELDLISWLCFSVL 700

RESULT 2
 AAB37797
 ID AAB37797 standard; Protein; 700 AA.

AC AAB37797;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Human interleukin-1 receptor intracellular ligand protein #4.
 XX
 KW Human; interleukin-1; IL-1; IL-1alpha; IL-1beta; IL-1 receptor;
 KW antiinflammatory; haemostatic; antibacterial; immunosuppressive;
 KW immunomodulatory; cardiac; cytostatic; neuroprotective; respiratory;
 KW inflammation; infection; sepsis; cachexia; autoimmune disorder;
 KW cardiovascular disorder; chronic myelogenous leukaemia;
 KW multiple sclerosis; inflammatory bowel disease; Crohn's disease.
 XX
 OS Homo sapiens.
 XX
 PN W0200064479-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 26-APR-2000; 2000MO-US11700.
 XX
 PR 27-APR-1999; 99US-0301274.
 XX
 PA (ANTI-) ANTIBODY SYSTEMS INC.
 XX
 PI Fredeking TM, Ignatyev GM;
 XX
 DR WPI; 2000-679646/66.
 XX
 PT Novel compositions comprising tetracycline or tetracycline-like
 PT compounds for the treatment and/or prevention of acute inflammatory
 PT responses and diseases, e.g. septic shock and immune complex-induced
 PT colitis -
 XX
 PS Disclosure; Page 159-162; 183pp; English.
 XX
 CC The present sequence is given in a specification relating to novel
 CC compositions and methods containing tetracycline or tetracycline-like
 CC compounds for treating and/or preventing acute inflammatory responses and
 CC diseases. Such diseases include acute inflammatory conditions associated
 CC with viral haemorrhagic diseases (including diseases caused by
 CC Bunyaviridae, Flaviviridae or Arenaviridae viruses),
 CC parasitic diseases, bacterial infections, sepsis, cachexia, autoimmune
 CC disorders, acute cardiovascular events, chronic myelogenous leukaemia and
 CC transplanted bone marrow-induced graft-versus-host disease, septic shock,
 CC immune complex-induced colitis, cerebrospinal fluid inflammation,
 CC multiple sclerosis, inflammatory responses associated with trauma,
 CC systemic inflammatory response syndrome (SIRS), adult respiratory
 CC distress syndrome (ARDS), acute liver failure, inflammatory bowel disease
 CC and Crohn's disease.
 XX
 SO Sequence 700 AA;
 Query Match 100.0%; Score 3713; DB 21; Length 700;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGIAAKLAKDREAEGLSHERAIKYNDYALNRECEAGTLFODSPFPAIPSLGF 60
 DB 1 MAGIAAKLAKDREAEGLSHERAIKYNDYALNRECEAGTLFODSPFPAIPSLGF 60
 QY 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAISLTNBEI 120
 DB 61 KELGPYSSKTRGMRKRPTEICADPOFIIGATRTDICOALGDCWLLAIAISLTNBEI 120
 QY 121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTXGELLFVHSAEGSEFWAL 180
 DB 121 LARVPLNOSFOENYAGIFHFQFQYGEWVVDRLPTXGELLFVHSAEGSEFWAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKRPPLFKIIQKALQKSLGCG 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEFTGIAEWELKRPPLFKIIQKALQKSLGCG 240

QY 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVSNGLQKILIRNPMGEVMTGWMND 300
 DB 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVSNGLQKILIRNPMGEVMTGWMND 300
 QY 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFSDLHYSLRLEICNLTPTLTSDTYKKMKLTK 360
 DB 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFSDLHYSLRLEICNLTPTLTSDTYKKMKLTK 360
 QY 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDEBDESGCTFLVGLIQKRRROR 420
 DB 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDEBDESGCTFLVGLIQKRRROR 420
 QY 421 KMGEDMHTTIGFGIYEVEPEELSGQTNHLSKNPFLTNRABRSPTFLNREVLRPKLPG 480
 DB 421 KMGEDMHTTIGFGIYEVEPEELSGQTNHLSKNPFLTNRABRSPTFLNREVLRPKLPG 480
 QY 481 EYLLVSTPEPNNDGPFCLRVFSEKADYQAVDDELEANLEEDDIDDGVRRLFAQ 540
 DB 481 EYLLVSTPEPNNDGPFCLRVFSEKADYQAVDDELEANLEEDDIDDGVRRLFAQ 540
 QY 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGKKEFYIL 600
 DB 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGKKEFYIL 600
 QY 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPQLHQVIYARPADQOLIIDFD 660
 DB 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPQLHQVIYARPADQOLIIDFD 660
 QY 661 NFVRCIVRLLETFLKIFKQDPENTGTIELDLISMLCFSVL 700
 DB 661 NFVRCIVRLLETFLKIFKQDPENTGTIELDLISMLCFSVL 700

RESULT 3

AAE25059 standard; Protein; 700 AA.

XX AAE25059;
 XX 30-OCT-2002 (first entry)
 XX Human calpain protein #2.
 XX Human; calpain; nervous system disorder; amyotrophic lateral sclerosis;
 KW Parkinson's disease; dementia; genito-urinary system disorder; stroke;
 KW Alzheimer's disease; multiple sclerosis; benign prostatic hyperplasia;
 KW urinary incontinence; gene therapy; cyostatic; nootropic; uropathic;
 KW neuroprotective.

XX Homo sapiens.
 XX PN W0200248326-A2.

XX 20-JUN-2002.

XX 14-DEC-2001; 2001WO-EP14819.

XX 14-DEC-2000; 2000US-255058P.

XX (FARB) BAYER AG.

XX Ramakrishnan S;

XX WPI; 2002-537625/57.

XX New human calpain polypeptide, useful for treating peripheral and
 PT central nervous system disorder and genito-urinary system disorders
 PT including urinary incontinence and benign prostatic hyperplasia -
 XX Disclosure; Page 101-104; 110pp; English.

XX The invention relates to novel human calpain proteins and polynucleotides
 CC encoding such proteins. Calpain sequences of the invention are useful for

CC treating, ameliorating or correcting dysfunctions or diseases such as
 CC peripheral or central nervous system (CNS) disorders (e.g., Parkinson's
 CC disease, Alzheimer's disease, multiple sclerosis, stroke, amyotrophic
 CC lateral sclerosis, dementia) and genito-urinary system disorders such
 CC as urinary incontinence and benign prostatic hyperplasia. They are also
 CC used in gene therapy. The present sequence is human calpain protein.

XX Sequence 700 AA;

Query Match 100.0%; Score 3713; DB 23; Length 700;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGIAAKLAKDRREAGLGSHERAIIKYLQDYEAALNECELEAGTLFPQDSFPAPALGF 60
 DB 1 MGIAAKLAKDRREAGLGSHERAIIKYLQDYEAALNECELEAGTLFPQDSFPAPALGF 60
 QY 61 KELGPYSSKTRGMRWRPTEICADPOFITGATRTDICGALGDCWLLAIAISLTINEET 120
 DB 61 KELGPYSSKTRGMRWRPTEICADPOFITGATRTDICGALGDCWLLAIAISLTINEET 120
 QY 121 LARVVPNLNOSFOENVAGIFHPQWQYGEWEVVDRLPTKQGLLFVHSABGSEFMSAL 180
 DB 121 LARVVPNLNOSFOENVAGIFHPQWQYGEWEVVDRLPTKQGLLFVHSABGSEFMSAL 180
 QY 181 LEKAVAKINGCYEALSGATTBEGEDFTGIAEWYELAKTPPFLFKIIOKALOKSLLGC 240
 DB 181 LEKAVAKINGCYEALSGATTBEGEDFTGIAEWYELAKTPPFLFKIIOKALOKSLLGC 240
 QY 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVSNGLQKILIRNPMGEVMTGWMND 300
 DB 241 SIDITSAADSEAITFOKLKVGHAYSVTGAEEVSNGLQKILIRNPMGEVMTGWMND 300
 QY 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFSDLHYSLRLEICNLTPTLTSDTYKKMKLTK 360
 DB 301 CPSEWNTIDPEERRLTRRHEDGFWMFSFSDLHYSLRLEICNLTPTLTSDTYKKMKLTK 360
 QY 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDEBDESGCTFLVGLIQKRRROR 420
 DB 361 MDGNMRGSTAGGCRNYPNTFMNPPOLYLKLEBEDDEDEBDESGCTFLVGLIQKRRROR 420
 QY 421 KMGEDMHTTIGFGIYEVEPEELSGQTNHLSKNPFLTNRABRSPTFLNREVLRPKLPG 480
 DB 421 KMGEDMHTTIGFGIYEVEPEELSGQTNHLSKNPFLTNRABRSPTFLNREVLRPKLPG 480
 QY 481 EYLLVSTPEPNNDGPFCLRVFSEKADYQAVDDELEANLEEDDIDDGVRRLFAQ 540
 DB 481 EYLLVSTPEPNNDGPFCLRVFSEKADYQAVDDELEANLEEDDIDDGVRRLFAQ 540
 QY 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGKKEFYIL 600
 DB 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIETCKIMVMDLSDSGSLGKKEFYIL 600
 QY 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPQLHQVIYARPADQOLIIDFD 660
 DB 601 WTKIQYOKIYREIDVRSGMTNSYEMRKALBEAGFMPQLHQVIYARPADQOLIIDFD 660
 QY 661 NFVRCIVRLLETFLKIFKQDPENTGTIELDLISMLCFSVL 700
 DB 661 NFVRCIVRLLETFLKIFKQDPENTGTIELDLISMLCFSVL 700

RESULT 4

AAB86128 standard; protein; 700 AA.

XX AAB86128;

XX 27-JUL-2001 (first entry)

XX Human calpain 80kDa subdomain protein fragment.

XX Calpain; calcium-activated cysteine proteinase; human; spatial structure;

KM Ca-activated cysteine proteinase; protein coordinate data; treatment;
KM structure-function study; ischemic condition; muscular dystrophy; tumor;
KM muscular; antitumor.
XX
OS Homo sapiens.
XX
PN EPI108779-A2.
XX
PD 20-JUN-2001.
XX
PF 13-DEC-2000; 2000EP-0127369.
XX
PR 14-DEC-1999; 99DE-1060225.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (PROT-) PROTEROS BIOSTRUCTURES GMBH.
XX
PI Strobl S, Fernandez-Catalan C, Bode W, Huber R, Suzuki K;
XX
XX WPI; 2001-376928/40.
XX
PT Spatial structures containing calpain-derived polypeptides, useful for
PT identifying calpain modulators and substrates, potentially useful e.g.
XX as antitumor agents -
XX
XX Claim 15; Fig 4; 182pp; German.
XX
CC This invention describes the novel spatial structure of human and rat
CC neutral calcium-activated cysteine protease (calpain) family. The spatial
CC structure (especially crystalline forms) are used for structure-function
CC studies, particularly for identifying (pseudo) substrates, inhibitors and
CC activators of calpains, potentially useful for treatment of ischemic
CC conditions, muscular dystrophy and/or tumors. The products of the
CC invention have anti-ischemic, muscular and antitumor activity. This
CC sequence represents the human calpain 80kDa subunit described in the
CC method of the invention.
XX
SQ Sequence 700 AA;

Query Match 99.5%; Score 3694; DB 22; Length 700;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 696; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAGIAAKLAKDREAEGLSHERAIKYLNDYEAALNCEBAGTLFQDPSPFPAISALGF 60
DB 1 MAGIAAKLAKDREAEGLSHERAIKYLNDYEAALNCEBAGTLFQDPSPFPAISALGF 60
QY 61 KELGPYSKTRGMWRKRPTEICADPQITIGATRTDTCGALGDCWLAIAIASLTNEEI 120
DB 61 KELGPYSKTRGMWRKRPTEICADPQITIGATRTDTCGALGDCWLAIAIASLTNEEI 120
QY 121 LARVVPINOSFOENYAGIFHFQFOWYGEWVWVDDRLPTDQGLLPVHSAEGSEFMSAL 180
DB 121 LARVVPINOSFOENYAGIFHFQFOWYGEWVWVDDRLPTDQGLLPVHSAEGSEFMSAL 180
QY 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
DB 181 LEKAVAKINGCYEALSGATTEGFEFTGIAEWELKPPNLFKIIQKALQKSLGCG 240
QY 241 SIDITSADSEALTFQKLVKHAIVTGAEEVNSGSIQKILIRINPGEVETGRMND 300
DB 241 SIDITSADSEALTFQKLVKHAIVTGAEEVNSGSIQKILIRINPGEVETGRMND 300
QY 301 CPSNNTTIDPEERELTRRHEDGEFMSFSDPLRHYSRLICNLTPDITLSTYKKKLT 360
DB 301 CPSNNTTIDPEERELTRRHEDGEFMSFSDPLRHYSRLICNLTPDITLSTYKKKLT 360
QY 361 MDGWRKRGSTAGCGENYPTFMANPOYLKLEEDDEDEDESGCTFLVGLIOKRRROR 420
DB 361 MDGWRKRGSTAGCGENYPTFMANPOYLKLEEDDEDEDESGCTFLVGLIOKRRROR 420
QY 421 KMGEDMTTIGFIVPEELSGQNNILSKNFFLTNRARSDPTFINREVANPKLP 480
DB 421 KMGEDMTTIGFIVPEELSGQNNILSKNFFLTNRARSDPTFINREVANPKLP 480

DB 421 KMGEDMTTIGFIVPEELSGQNNILSKNFFLTNRARSDPTFINREVANPKLP 480
QY 481 EYILVPESTPEPNKDGDFCIRVFSEKKADYQAVODEIANLEPDISDIDDGVRRLFAQ 540
DB 481 EYILVPESTPEPNKDGDFCIRVFSEKKADYQAVODEIANLEPDISDIDDGVRRLFAQ 540
QY 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIECTKIWMDLSDSGSKGLKEFYIL 600
DB 541 LAGEDAEISAFELQTLIRVLAARODIKSDGFSIECTKIWMDLSDSGSKGLKEFYIL 600
QY 601 WTKIQKQKRYREIDVDRSGTMSYEMRKALBEAGFMPQQLHOVYARPADQOLIIPD 660
DB 601 WTKIQKQKRYREIDVDRSGTMSYEMRKALBEAGFMPQQLHOVYARPADQOLIIPD 660
QY 661 NFVRCVRLTETLFKIFKQIDPENTGIIEDLISMLGFSYL 700
DB 661 NFVRCVRLTETLFKIFKQIDPENTGIIEDLISMLGFSYL 700

RESULT 5
AAB86130
ID AAB86130 standard; protein; 700 AA.
XX
AC AAB86130;
XX
DT 27-JUL-2001 (first entry)
XX
DE Rat calpain 80kDa subdomain protein fragment.
XX
KM Calpain; calcium-activated cysteine proteinase; human; spatial structure;
KM Ca-activated cysteine proteinase; protein coordinate data; treatment;
KM structure-function study; ischemic condition; muscular dystrophy; tumor;
KM muscular; antitumor.
XX
OS Rattus norvegicus.
XX
PN EPI108779-A2.
XX
PD 20-JUN-2001.
XX
PF 13-DEC-2000; 2000EP-0127369.
XX
PR 14-DEC-1999; 99DE-1060225.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (PROT-) PROTEROS BIOSTRUCTURES GMBH.
XX
PI Strobl S, Fernandez-Catalan C, Bode W, Huber R, Suzuki K;
XX
XX WPI; 2001-376928/40.
XX
PT Spatial structures containing calpain-derived polypeptides, useful for
PT identifying calpain modulators and substrates, potentially useful e.g.
XX as antitumor agents -
XX
XX Claim 15; Fig 6; 182pp; German.
XX
CC This invention describes the novel spatial structure of human and rat
CC neutral calcium-activated cysteine protease (calpain) family. The spatial
CC structure (especially crystalline forms) are used for structure-function
CC studies, particularly for identifying (pseudo) substrates, inhibitors and
CC activators of calpains, potentially useful for treatment of ischemic
CC conditions, muscular dystrophy and/or tumors. The products of the
CC invention have anti-ischemic, muscular and antitumor activity. This
CC sequence represents the rat calpain 80kDa subunit described in the
CC method of the invention.
XX
SQ Sequence 700 AA;

Query Match 94.7%; Score 3518; DB 22; Length 700;
Best Local Similarity 93.4%; Pred. No. 7.7e-310;
Matches 654; Conservative 26; Mismatches 20; Indels 0; Gaps 0;

```

QY 1 MAGIATLADREAEGLGSHERAIKYLNDYELRNECEAGTLFODPSFPAISALGF 60
DB 1 MAGIATLADREAEGLGSHERAIKYLNDYELRNECEAGTLFODPSFPAISALGF 60
QY 61 KELGPYSSKTRGMKWKPTREICADPOITIGGATTTDCCGALGDCWLLAIASTLTNEEI 120
DB 61 KELGPYSSKTRGMKWKPTREICADPOITIGGATTTDCCGALGDCWLLAIASTLTNEEI 120
QY 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFMSAL 180
DB 121 LARVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFMSAL 180
QY 181 LERAYAKINGCYEALSGATTEGPDPTGGIAEWELKRPENLFTKIQALOKGSLIGC 240
DB 181 LERAYAKINGCYEALSGATTEGPDPTGGIAEWELKRPENLFTKIQALOKGSLIGC 240
QY 241 SIIITSAADSEALTFQGLVGHAYSVTGAEEVSSGLQKILIRNPMGQVEMTGKWNND 300
DB 241 SIIITSAADSEALTFQGLVGHAYSVTGAEEVSSGLQKILIRNPMGQVEMTGKWNND 300
QY 301 CPSPMTIDPEERERLTERHEDGEFMSFSDFLRHYSRLICNLTPDTLTSDTYKMWLTK 360
DB 301 CPSPMTIDPEERERLTERHEDGEFMSFSDFLRHYSRLICNLTPDTLTSDTYKMWLTK 360
QY 361 MDGWRGSGTAGGCRNYPNTFMNMPQYLKLEEDDEDEDEGSCCTFLVGLIQHRRROR 420
DB 361 MDGWRGSGTAGGCRNYPNTFMNMPQYLKLEEDDEDEDEGSCCTFLVGLIQHRRROR 420
QY 421 KMGEDMTTIGFGEVEEELSGOTNHLKSNFELTNARSRDPTNLRVLRNFKLP 480
DB 421 KMGEDMTTIGFGEVEEELSGOTNHLKSNFELTNARSRDPTNLRVLRNFKLP 480
QY 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEEDDIDGVRRLFAQ 540
DB 481 EYLLVSTFEPNKGDFCIRVFSEKADYQAVDEIEANLEEDDIDGVRRLFAQ 540
QY 541 LAGBDAISAFELQTIIRVLAARQDIKSDGFSIETCKIWMMLSDSGSLGKEFYIL 600
DB 541 LAGBDAISAFELQTIIRVLAARQDIKSDGFSIETCKIWMMLSDSGSLGKEFYIL 600
QY 601 WTKIQOKYKTYRELDVDSGTMSYEMRKALEEGFMPQOLHOVYARADQOLIIDPD 660
DB 601 WTKIQOKYKTYRELDVDSGTMSYEMRKALEEGFMPQOLHOVYARADQOLIIDPD 660
QY 661 NFVRCVRLTELFKIFKQDPENTGTLELDLSWLCFSVL 700
DB 661 NFVRCVRLTELFKIFKQDPENTGTLELDLSWLCFSVL 700

```

```

RESULT 6
AAE23085
ID AAE23085 standard; Protein: 713 AA.
XX
AC AAE23085;
XX
DT 21-AUG-2002 (first entry)
XX
DE Calcium-activated neutral protease protein.
XX
KW Transgenic; transgenic animal; pharmacological therapy; gene therapy;
KW phenotype modulation; calcium-activated neutral protease; enzyme;
KW genetic disease.
XX
OS Unidentified.
XX
PN MO200203787-A2.
PD 17-JAN-2002.
XX
PF 06-JUL-2001; 2001WO-US21427.
XX
PR 06-JUL-2000; 2000US-216109P.
PR 06-JUL-2000; 2000US-216251P.

```

```

PR 06-JUL-2000; 2000US-216258P.
PR 06-JUL-2000; 2000US-216768P.
PR 10-JUL-2000; 2000US-217449P.
PR 10-JUL-2000; 2000US-217450P.
PR 10-JUL-2000; 2000US-217650P.
PR 27-JUL-2000; 2000US-221491P.
PR 27-JUL-2000; 2000US-221659P.
PR 27-JUL-2000; 2000US-221670P.
PR 07-AUG-2000; 2000US-223170P.
PR 07-AUG-2000; 2000US-223172P.
PR 07-AUG-2000; 2000US-223460P.
PR 26-OCT-2000; 2000US-244037P.
PR 26-OCT-2000; 2000US-244111P.
PR 26-JUN-2001; 2001US-301217P.
XX
PA (DELTA-) DELTAGEN INC.
XX
PI Allen KD, Leviten MW;
XX
DR WPI, 2002-154853/20.
DR N-PSDB; AAD37051.
XX
PT Novel non-human transgenic animal, preferably transgenic mice
PT comprising disruption in target gene, e.g., trypsin gene, useful for
PT identifying an agent that modulates expression or function of target
PT gene.
XX
PS Disclosure; Fig 14; 74pp; English.
XX
CC The present invention relates to non-human transgenic animals preferably
CC transgenic mice comprising disruption in target gene such as trypsin
CC gene. The invention also relates to compositions and methods relating
CC to the characterisation of gene functions. The transgenic animals are
CC useful for identifying an agent that modulates the expression or function
CC of a target. They are useful for identifying an agent that modulates a
CC phenotype associated with a disruption in trypsin gene or limulus
CC clotting factor protease-like genes by administering an agent to the
CC transgenic animal and determining whether the agent modulates the
CC phenotype where the agent has effect on decreased body weight, decreased
CC thymus weight, decreased thymus to body weight ratio, increased pre-pulse
CC inhibition, significant decrease in their response latency to the hot
CC plate test or a decreased response threshold to metrazol. Agents that
CC modulate the expression, function or activity of the target gene are
CC useful for treating a disorder associated with a mutation in trypsin
CC gene or in limulus clotting factor protease-like gene. The transgenic
CC animals are useful for testing the efficacy of proposed genetic and
CC pharmacological therapies for human genetic diseases. They are useful
CC as models for diseases, disorders or conditions associated with
CC phenotypes relating to a disruption in a target and to identify drugs,
CC pharmaceuticals, therapies and interventions which may be effective in
CC treating a disease or other phenotypic characteristics of the animal.
CC The present sequence is calcium-activated neutral protease protein.
CC This sequence is used in the exemplification of the invention.
XX
SQ Sequence 713 AA;

```

```

Query Match 65.6%; Score 2435; DB 23; Length 713;
Best Local Similarity 62.8%; Pred. No. 1.6e-211; Mismatches 4; Gaps 3;
Matches 437; Conservative 126; Mismatches 129; Indels 4; Gaps 3;

```

```

QY 3 GIAAKLAKREAEAGLGSHERAIKYLNDYELRNECEAGTLFODPSFPAISALGF 62
DB 13 GVAQAQYOKKRDKEGLGRHENAIIKYLQGVETLRARCLQSGVLFQDEAPRPVSHSGFKE 72
QY 63 LEPYSSKTRGMKWKPTREICADPOITIGGATTTDCCGALGDCWLLAIASTLTNEEI 122
DB 73 LGPHSSKTYGKWKRPTELSMNPQFTVDGATRTDCCGALGDCWLLAIASTLTNEEI 122
QY 123 RVPVPLNOSFOENYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFMSAL 182
DB 133 RVPVYQSGFQDGYAGIFHFQFQWYGEWVEVVDDRLPTKDGELLFVHSAEGSEFMSAL 192
QY 183 KAYAKINGCYEALSGATTEGPDPTGGIAEWELKRPENLFTKIQALOKGSLIGC 242

```

```

Db      193 KAYKVNQSYALSGCTSEAFEDFTGVTWMDLQKAPSLYQIILALRSGSLGCSI 252
Qy      243 DITAADEATPFOKLIVKGAHVSVTGAEVESNGSLQILIRNPMGEVMTGRNDPCP 302
Db      253 NISIRDLKATLTKLVKGVGHVSTGAKQVYQOGKRVNLIRNPMGEVMTGRNDPCP 312
Qy      303 SWNTIDPEERERLIRRHEDGFWMSFSDPLHYSRLICNLTPDTLSDYKKWKLTKMD 362
Db      313 EMNVKDPYERBQLRVKMGDEGFWMSFRDPIREFTLLEICNLTPDLKSRILRNMTTPE 372
Qy      363 GNMRRGTSAGCCRYNPTFMNPOYLKLEED--EDEDESGCTFLVGLIQKRRQR 420
Db      373 GWMRGSTAGCCRYNPTFMNPOYLKLEED--EDEDESGCTFLVGLIQKRRQR 432
Qy      421 KMGDMHTIGIYEVPEELSGQTNILHSKNFPLTNRARESDPILNREVLANFKLPBG 480
Db      433 RFGDMETIGPAVVOVPEELAGQP-VHLKRPFLANASRAOSEHFINREVSNRLRPPG 491
Qy      481 EYLVPTFEENKGDGFCIRVSEKKADYQAVDEIEANL--EEDPISDDIDGVRRLFA 539
Db      492 EYLVPTFEENKGDGFCIRVSEKKADYQAVDEIEANL--EEDPISDDIDGVRRLFA 551
Qy      540 QLAGEDAFISAFELQTLIRRLAKRQDKSGFSIETCKIWMMLDSGSKGLKERYI 599
Db      552 KLAGDMETISVKEIQTILNRIISRKQDRTNGFSLSCRSWNLMDRDGNGKLGIVENI 611
Qy      600 LMTIOLKQKQIYREIDVDKSGTMSYEMRKALEEAGFPCQLHGVIAVARPADQLIDF 659
Db      612 LMRIRNLYTLIRKFDLDSGSMAYEMRKALIEAGFKLANKLHELIITRISBDLAVDF 671
Qy      660 DNFRCLVRLTLPFKIFKQDLPENTGTELDLISWL 695
Db      672 DNFRCLVRLTLPFKIFKQDLPENTGTELDLISWL 707

RESULT 7
AAW41564
ID AAW41564 standard; Protein; 703 AA.
XX
AC AAW41564;
XX
DT 27-APR-1998 (first entry)
XX
DE Human calpain.
XX
KW Calpain; human; leukocyte; calcium dependent cysteine protease;
KW screening; activator; inhibitor; treatment; prevention; cancer;
KW cerebral apoplexy; cerebral infarction; subarachnoid haemorrhage;
KW Alzheimer's disease; myodystrophy; cataracts; collagen disease;
KW ischaemic heart disease; atherosclerosis; arthritis.
XX
OS Homo sapiens.
XX
EN EP799892-A2.
XX
PD 08-OCT-1997.
XX
PF 03-APR-1997; 97EP-010508.
XX
PR 05-APR-1996; 96JP-0083649.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Kawamoto T, Nishi K, Shintani Y.
XX
DR WPI, 1997-482674/45.
XX
DR N-PSDB; AAV04202.
XX
PT Human calpain protein and related DNA - useful for drug screening
XX and treating cancer, stroke, etc.
PS Claim 1; Pages 23-25; 43pp; English.

```

```

XX      The present sequence is calpain, a human leukocyte derived
CC      calcium dependent cysteine protease. Calpain can be used to screen
CC      for compounds that activate or inhibit its proteolytic activity.
CC      Calpain DNA can be used to treat or prevent cancer, cerebral
CC      apoplexy, cerebral infarction, subarachnoid haemorrhage,
CC      Alzheimer's disease, myodystrophy, cataracts, ischaemic heart
CC      disease, atherosclerosis, arthritis or collagen disease.
XX
SQ      Sequence 703 AA;
Query Match 63.8%; Score 2370.5; DB 18; Length 703;
Best Local Similarity 61.5%; Pred. No. 1.1e-205;
Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;
Qy      1 MAGIAALADREAEGISHERAIKYNQVLEALRECEAGTLFODPFPPIPSALGF 60
Db      1 MAQAQAVSQKQANATQGLSNNBLKTLGQDFTLKQCCIDSGVLFKDPFPFACPSALGY 60
Qy      61 KELAPVSSKTRGMRKRPTEICADPOFIIGATRTDICGALDGCWLLAIASTLTNEE 120
Db      61 KDLGPGSPQIGIIMKRPTELCPSPQITVGATRTDICGALDGCWLLAIASTLTNEE 120
Qy      121 LARVPLNQSFOENYAGIFHFQWQYGEWVYVDRLPRKQDELFLVHABSGSEFSA 180
Db      121 LVRVPRDQFOENYAGIFHFQWQYGEWVYVDRLPRKQQLFLHSEQNEFWSAL 180
Qy      181 LEKAVKINQCYEALSGATTEGPEPTGIAEYELKRPPIFKIIOKALQKSLIGC 240
Db      181 LEKAVKINQCYEALSGATTEGPEPTGIAEYELKRPPIFKIIOKALQKSLIGC 240
Qy      241 SIDTSSADEATPFOKLIVKGAHVSVTGAEVESNGSLQILIRNPMGEVMTGRNDPCP 300
Db      241 SIDVSSAAEAEATISQGLVSHAVSVTGAEVNFQGHPELILIRNPMGEVMTGRNDPCP 300
Qy      301 CPSWNTIDPEERERLIRRHEDGFWMSFSDPLHYSRLICNLTPDTLSDYKKWKLTK 360
Db      301 APEWNHIDPPRKKELDKVDGGEFWMSLSDFFVQFSLRLICNLSPDLSSEVHKMNLVL 360
Qy      361 MDGNMRRGTSAGCCRYNPTFMNPOYLKLEEDDEED--GESGCTFLVGLIQKRRR 418
Db      361 FNGHMTGTSAGCCRYNPTFMNPOYLKLEEDDEED--GESGCTFLVGLIQKRRR 420
Qy      419 QKMGEDMTIGIYEVPEELSGQTNILHSKNFPLTNRARESDPILNREVLANFKLP 478
Db      421 RKRIGQGLSIGAVVQVPELSEHTDAHLGRDFFLAQGSARSTVYVNRVSGRARLP 480
Qy      479 PGEYLVPTFEENKGDGFCIRVSEKKADYQAVDEIEANLBEFDISE--DDIDGVRRL 537
Db      481 PGEYLVPTFEENKGDGFCIRVSEKKADYQAVDEIEANLBEFDISE--DDIDGVRRL 540
Qy      538 PAOLAGEDAISAFELQTLIRRLAKRQDKSGFSIETCKIWMMLDSGSKGLKERYI 597
Db      541 FEKLAKGDSITANALKILNLAEPKRTDIKFGDFNINTREMI SLDSNGTGLGAVEF 600
Qy      598 YIIMTKIQKQIYREIDVDKSGTMSYEMRKALEEAGFPCQLHGVIAVARPADQLID 657
Db      601 KTLMLKIQKQIYREIDVDKSGTMSYEMRKALEEAGFPCQLHGVIAVARPADQLID 660
Qy      658 DFDNFVRCVRLTLPFKIFKQDLPENTGTELDLISWL 696
Db      661 NFDSEFVACMTRLETLPFLSLDDEKDGWQSLAEWLC 699

RESULT 8
AAW41565
ID AAW41565 standard; Protein; 712 AA.
XX
AC AAW41565;
XX
DT 27-APR-1998 (first entry)
XX
DE Human calpain.

```

XX Calpain; human; leukocyte; calcium dependent cysteine protease;
 KW screening; activator; inhibitor; treatment; prevention; cancer;
 KW cerebral apoplexy; cerebral infarction; subarachnoid haemorrhage;
 KW Alzheimer's disease; myodystrophy; cataracts; collagen disease;
 KW ischemic heart disease; atherosclerosis; arthritis.
 OS Homo sapiens.
 PN EP799892-A2.
 XX
 PD 08-OCT-1997.
 XX
 PF 03-APR-1997; 97EP-0105508.
 XX
 PR 05-APR-1996; 96JP-0083649.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Kawamoto T, Nishi K, Shintani Y;
 DR WPI; 1997-482674/45.
 DR N-PSDB; AAV04203.
 XX
 PT Human calpain protein and related DNA - useful for drug screening
 PT and treating cancer, stroke, etc.
 XX
 PS Claim 2; Pages 26-28; 43pp; English.
 CC The present sequence is calpain, a human leukocyte derived
 CC calcium dependent cysteine protease. Calpain can be used to screen
 CC for compounds that activate or inhibit its proteolytic activity.
 CC Calpain DNA can be used to treat or prevent cancer, cerebral
 CC apoplexy, cerebral infarction, subarachnoid haemorrhage,
 CC Alzheimer's disease, myodystrophy, cataracts, ischaemic heart
 CC disease, atherosclerosis, arthritis or collagen disease.
 XX
 SQ Sequence 712 AA;
 Query Match 63.8%; Score 2370.5; DB 18; Length 712;
 Best Local Similarity 61.5%; Pred. No. 1.1e-205;
 Matches 430; Conservative 123; Mismatches 143; Indels 3; Gaps 2;

QY 1 MAGIAKLAADREAEAGLSHERAIKYNODYALNRECEAGTLFQDPSPAPISALGF 60
 DB 10 MAQAAGVSRQRAATGSGNQAALKYLGGDFKTLRQCCLDSGLVLPDPPFACPSALGY 69
 QY 61 KEIGPSSKTRGRKRPTEICADPPTIGATRTDICGALGDCWLLAIAISLTINEET 120
 DB 70 KDLGPGSPOTGGIIMKRPTELCPSPQPIVGATRTDICGALGDCWLLAIAISLTINEEL 129
 QY 121 LARVPLNOSFOENVAGIFHFOFMOYGEWVVDRLPTKDELLFVHSAEGSEFMSAL 180
 DB 130 LRVVPRDDPQENYAGIFHFOFMOYGEWVVDRLPTKDELLFVHSEQNEFMSAL 189
 QY 181 LEKAYAKINGCYALSGATTEGFEDEFTGIAEYELKPPNLFKIIQALQKSLAGC 240
 DB 190 LEKAYAKINGCYALSGATTEGFEDEFTGIAEYELKPPNLFKIIQALQKSLAGC 249
 QY 241 SIDIITAADESEAITFOQLVKGHAYSTGAEEVNSGLQTLIRINWGEVENTGRMND 300
 DB 250 SIVSSAAEAETISOKLVKSHAYSTGVEVNFQGHPEKILIRINWGEVENSGLMSDD 309
 QY 301 CPSMNTIDPEERELIRRHEDGEFMSFDELHYSLETCNLTPTLTSDTYKKMKLTK 360
 DB 310 APENHIDPRKEELDKVDEGEFMSLDFVQFSLTECNLTPTLTSDTYKKMKLTK 369
 QY 361 MDGNMRSSTAGCCRNYPNTFMNPOYLKLEEDDEDEED--GESGCTFVGLIQRKRR 418
 DB 370 FNGHMTGRSTAGCCRNYPNTFMNPOYLKLEEDDEDEED--GESGCTFVGLIQRKRR 429
 QY 419 QRKMGEDMTITGIGIYVEPEELSGQNIHLKSKFELTNRAKRSDFFINIREVLNRPKL 478

DB 430 RKRIGQMGSLIGYAVYQVPEKEHSHTDAHLGRDFLAPPSARTSTYVNLREVSGARLP 489
 QY 479 PGEYIIVPSTFPEBNKGDGFCIRVSEKKADYQAVDEIBANLEEFPISE-DDIDGVRRL 537
 DB 490 PGEYIIVPSTFPEBNKGDGFCIRVSEKKADYQAVDEIBANLEEFPISE-DDIDGVRRL 549
 QY 538 PAQLAGEBAEISAFELQTLIRRYLARNRQDIKSDGFSIETCKIWMMLDSDGSKGLKEF 597
 DB 550 FEKLAGEDESEITANALKILNEAFSKRTDIKFGDFNINTCREWISLSDSGTGLGAVEF 609
 QY 598 YIIMTKIQKQKRYREIDVDRSGTMSYERKALBEAGFPMPCQLHQVYVARADQOLI 657
 DB 610 KTLMLKIQKQKRYREIDVDRSGTMSYERKALBEAGFPMPCQLHQVYVARADQOLI 669
 QY 658 DFDNFVRLVRLTEFLFKIFKQDPDENTGTIEDLISMLC 696
 DB 670 NFDSPVACMIRLETFLKLSLDEDDQGVQSLAEMLC 708

RESULT 9
 AAE14338
 ID AAE14338 standard; Protein; 703 AA.
 XX
 AC AAE14338;
 XX
 DT 07-MAR-2002 (first entry)
 XX
 DE Human protease PRN3-3 protein.
 XX
 KW Human; protease; PRN3-3; tranquilliser; gene therapy; vaccine; allergy;
 KW infection; dermatitis; atherosclerosis; rheumatoid arthritis; hepatitis;
 KW atherosclerosis; psoriasis; Alzheimer's disease; mental disorder; cancer;
 KW gastrointestinal disorder; Cushing's syndrome; seizure; glaucoma; stroke;
 KW epithelial disorder; urticaria; anorexia; trauma; asthma; eczema; nausea;
 KW hypertension; neurological disorder; Parkinson's disease; drug screening;
 KW cardiac; cell proliferative disorder; multiple sclerosis; osteoporosis;
 KW diabetes mellitus; glomerulonephritis; cardiovascular disorder; anaemia;
 KW autoimmune disorder; inflammatory disorder; myocardial infarction; AIDS;
 KW developmental disorder; reproductive disorder; infertility; diarrhoea;
 KW dementia; acidosis; cataract; gynaecomastia; epilepsy; jaundice.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..508
 FT Domain /note="Calpain catalytic domain"
 FT Domain 45..344
 FT Domain /note="Calpain family cysteine protease domain"
 FT Domain 99..110
 FT Domain /note="Thiol protease Cys motif"
 FT Domain 355..512
 FT Domain /note="Calpain larger subunit domain III"
 XX
 PN WO200183775-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 04-MAY-2001; 2001MO-US14651.
 XX
 PR 04-MAY-2000; 2000US-202082P.
 PR 11-MAY-2000; 2000US-203566P.
 PR 17-MAY-2000; 2000US-205803P.
 PR 25-MAY-2000; 2000US-207477P.
 PR 01-JUN-2000; 2000US-209402P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Deleage AM, Lal P, Hafalia A, Patterson C, Wallia NK, Kearney L;
 PI Tribouley CM, Khan FA, Yao MG, Baughn MR, Azimzai Y, Elliott VS;
 PI Nguyen DB, Gandhi AR, Yang J, Hernandez R, Policky UL, Lu DM;
 PI Reddy R, Yue H, Tang YJ;
 DR WPI; 2002-034518/04.

DR N-PSDB; AAD23844.
 XX Novel human proteases and polynucleotides encoding the proteases,
 PT useful for treating, diagnosing or preventing cell proliferative,
 PT cardiovascular, autoimmune/inflammatory, neurological and developmental
 PT disorders -
 PS
 PS Claim 1; Page 121-123; 151pp; English.
 XX
 CC The invention relates to human proteases (PRTS1-14) and its corresponding
 CC cDNA molecules. Human PRTS and its nucleic acid molecule are useful for
 CC the diagnosis, treatment and prevention of disorders associated with
 CC increased or decreased expression of PRTS. Examples of such disorders
 CC include, cell proliferative disorders (arteriosclerosis, atherosclerosis,
 CC hepatitis, psoriasis and cancers); autoimmune/inflammatory disorders
 CC (AIDS, Addison's disease, allergy, anaemia, asthma, atopic dermatitis,
 CC diabetes mellitus, glomerulonephritis, multiple sclerosis, osteoporosis,
 CC trauma, Grave's disease, rheumatoid arthritis, ulcerative colitis, and
 CC viral, bacterial, fungal, parasitic, protozoal and helminthic
 CC infections); cardiovascular disorders (myocardial infarction, ischaemic
 CC heart disease and hypertension); neurological disorders (epilepsy,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, dementia,
 CC Parkinson's disease, stroke, mental disorders including mood, anxiety
 CC and seasonal affective disorder and prion diseases); gastrointestinal
 CC disorders (Crohn's disease, anorexia, nausea, diarrhoea and jaundice);
 CC epithelial disorders (contact dermatitis, eczema, acne vulgaris,
 CC alopecia, scabies, insect bites and urticaria); reproductive disorder
 CC (infertility, disruption of estrous and menstrual cycle and
 CC gynaecomastia); and developmental disorders (renal tubular acidosis,
 CC Cushing's syndrome, seizure disorders, congenital glucocortic acid and cataract).
 CC PRTS DNA is also in useful in gene therapy. PRTS and its immunogenic
 CC fragments are useful for screening libraries of compounds in several drug
 CC screening assays. The present sequence is human protease PRTS-3 protein.
 XX
 SQ Sequence 703 AA;
 Query Match 63.8%; Score 2367.5; DB 23; Length 703;
 Best Local Similarity 61.5%; Pred. No. 2.1e-205;
 Matches 430; Conservative 122; Mismatches 144; Indels 3; Gaps 2;
 QY 1 MAGIAAKLAKREAAEGGSHERRAITYLNQDYEARNECEAGTLPDPPSPALPSALGF 60
 DB 1 MAAQAAAGVSRRQATQGGSNQNALKYIGDPFKTLRQCCDLSGVLFKDPPEPAPSPALGY 60
 QY 61 KELGPYSKTRGMKMRKPTETICADPOFIIGATRTDIOGALGDCWMLAALASLTINEI 120
 DB 61 KOLGPSPQTOGIITWKRTETLCPSPQIVGATRTDIOGALGDCWMLAALASLTINEI 120
 QY 121 LARVVPINQSFQENYAGIFHFQFQWQYGEVWVVDRLPTXQGLLFFVSAEGSEFWSAL 180
 DB 121 LYRVVPIDQFQENYAGIFHFQFQWQYGEVWVVDRLPTXQGLLFFVSAEGSEFWSAL 180
 QY 181 LEKAYAKINGCYEALSGATTEGFEPTFGTAEMWELKKPPNLFKTIQKALQKSLIGC 240
 DB 181 LEKAYAKINGCYEALSGATTEGFEPTFGTAEMWELKKPPNLFKTIQKALQKSLIGC 240
 QY 241 SIDDTSAADSEAFIFQKLVKGAHSVTCAGEVNSGQKILIRINPGEVEMTGRANDN 300
 DB 241 SIDDTSAADSEAFIFQKLVKGAHSVTCAGEVNSGQKILIRINPGEVEMTGRANDN 300
 QY 241 SIDVYSAAEALITISQKLVKSHAYSVTGVEEVNQGHKKILRLNPGVEVMSGANSDD 300
 DB 241 SIDVYSAAEALITISQKLVKSHAYSVTGVEEVNQGHKKILRLNPGVEVMSGANSDD 300
 QY 301 CPWNNTIDPEERERLIRRHEDGEFWMSPDFLRHYSRLICNLTPDTLSTDYKKMKLTK 360
 DB 301 CPWNNTIDPEERERLIRRHEDGEFWMSPDFLRHYSRLICNLTPDTLSTDYKKMKLTK 360
 QY 361 MDGWRKRGSTIGGCRNRYNTFMWNPQYILKEEBDEEDD--GESGCTFLVGLIQKRRR 418
 DB 361 MDGWRKRGSTIGGCRNRYNTFMWNPQYILKEEBDEEDD--GESGCTFLVGLIQKRRR 418
 QY 419 ORKMGEDHTIGFGIYEVBEELSGQTNHLKSNFLTRARAPRSTFNLREVNRFLPL 478
 DB 419 ORKMGEDHTIGFGIYEVBEELSGQTNHLKSNFLTRARAPRSTFNLREVNRFLPL 478
 QY 421 KRGIGQGLMISIGYAVYQVFKELBSHTDHLGRDFLVAQPARSTTYNLEVSGRARLP 480
 DB 421 KRGIGQGLMISIGYAVYQVFKELBSHTDHLGRDFLVAQPARSTTYNLEVSGRARLP 480
 QY 479 PGEYILVPTPEPNKDGFCIRVSEKKADYQAVDEIEANLIEEPDISE-DDIDGGVRL 537

DB 481 PGEYILVPTPEPNKDGFCIRVSEKKADYQAVDEIEANLIEEPDISE-DDIDGGVRL 540
 QY 538 FAQAGDADATSAFELQTIARRVLAKRQDIKSPGSIETCKIWMVDDSPGSKLGEKEF 597
 DB 541 FEKLAKGDSSTIANALKITLNEAPSKRTDKFDFGNINNTRENISLSDSGTGLGAVF 600
 QY 598 YIIMTKIQYQKTYREIDVDRSGTMNSYEMRKALIEAGFPCQLHQVIYAREPADQLI 657
 DB 601 KITLMKIQKLEIYEMWDVNHSGTIDAHMERKTLRRAGFLINSQVQOTIALRYACSKLGI 660
 QY 658 DFDNFVRCVLRLTELFKIFKQDPENTGTELDLISWLC 696
 DB 661 NFDSPVACMRILETLFELFLDDEKXQGWQVLSIAEWLC 699
 RESULT 10
 AAU72884
 ID AAU72884 standard; Protein; 703 AA.
 AC
 XX AAU72884;
 DT 26-FEB-2002 (first entry)
 DE
 XX Human aspartyl protease partial protein sequence #9.
 XX Human; protease; PCR primer; cytosolic; immunomodulator; cardiant;
 KW vasoactive; antidiarrhoeal; analgesic; endocrine; nootropic; tranquiliser;
 KW hypertensive; hypotensive; neuroleptic; neuroprotective; anabolic;
 KW anorectic; antiinflammatory; aspartyl protease; cysteine protease;
 KW metalloprotease; serine protease; cancer; haematopoietic; breast; colon;
 KW lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;
 KW immune-related disease; cardiovascular disease; neuronal disease;
 KW migraine; sexual dysfunction; mood disorder; attention disorder;
 KW cognition disorder; hypotension; hypertension; psychotic disorder;
 KW dyskinnesia; metabolic disorder; inflammatory disorder.
 KW
 XX Homo sapiens.
 OS
 XX WO200183782-A2.
 PN
 XX 08-NOV-2001.
 PD
 XX 04-MAY-2001; 2001WO-US14431.
 PF
 XX 04-MAY-2000; 2000US-201879P.
 PR
 XX (SUGEN-) SUGEN INC.
 PA
 XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
 PI Payne V;
 PL
 XX WPI, 2002-041502/05.
 DR N-PSDB; AAS97167.
 DR
 XX
 PT Novel protease polypeptide useful for screening for substances that may
 PT be used to treat, e.g., cancers, immune-related diseases,
 PT cardiovascular disease, migraine, pain, psychotic and inflammatory
 PT disorders -
 PS
 PS Claim 28; Figure 2B; 232pp; English.
 XX
 CC The invention relates to an isolated, enriched, or purified protease
 CC polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
 CC screen for substances (S) that may modulate its activity. Administering
 CC S (which modulates protease activity in vitro) may be used to treat a
 CC disease or disorder selected from cancers (e.g., of tissues, of blood or
 CC haematopoietic origin, of the breast, colon, lung, prostate, cervical,
 CC brain, ovarian, bladder or kidney), immune-related diseases and
 CC disorders, cardiovascular disease, brain or neuronal-associated diseases
 CC (e.g., central or peripheral nervous system diseases, migraine, pain,
 CC sexual dysfunction, mood disorders, attention disorders, cognition
 CC disorders, hypertension, hypertension, psychotic disorders, neurological

CC disorders and dyskinesias), metabolic disorders and inflammatory
CC disorders. (1) may also be useful as a diagnostic tool for a disease or
CC disorder such as those above. AAU2876-AAU2910 represent human
CC protease amino acid sequences of the invention.

SQ Sequence 703 AA;

Query Match	63.8%	Score 2367.5	DB 23	Length 703
Best Local Similarity	61.5%	Pred. No. 2.1e-205		
Matches 430	Conservative 122	Mismatches 144	Indels 3	Gaps 2

QY	1	MAGIAAKLAKDRAAAGAGSHRAIKYLNQDVEALRNLCLEAGTLFQDPSPPALPSALGF	60
Db	1	MAQQAAGSRRPAAITAGLSSNQNALKTYLGQDFKTLRQGLCSGVLFKQEPFPACPSALGY	60
QY	61	KELGPYSSKTRGRMWRPTEICADPOFLIGGATRTDICOAGLGDWCLAAIASLTNBEI	120
Db	61	KDLGPSPQQTQIGIMWRPELCPSPQFLVIGATRTDICOGLGDWCLAAIASLTNBEI	120
QY	121	LAAVPLNQSFOENAVGIFHFQFQWQYGEWVEVVDRLPTYKQDELLFVHSAESGEPMSAL	180
Db	121	LYVIVRRDDPFOENYAGIFHFQFQWQYGEWVEVVIDRLPTYKNGQLFLFHSQGNPEMSAL	180
QY	181	LEAAVAKINGCCVEALSGGATGTFEPFGTIGIAEWELKKRPNNPKITIOKLQGSILGC	240
Db	181	LEAAVAKINGCCVEALSGGATGTFEPFGTIGTGGISEFLDKKPPNNLQIIRKLCASSLGC	240
QY	241	SIDITSAADSEAITFOKLVKGHAVSYTGAEEVNSGSLQKTLRIENPMGEVEMTGRWMD	300
Db	241	SIDVYSAAEAEAITSGKLVKSHAVSYTGVEEVNPQGHPEKILRIENPMGEVEMSGAMSD	300
QY	301	CPBWNTIIDBERRLTRRHEDGPFMWSFSDFLRHYSLEICNLIPDTLTSPTTKKMLTK	360
Db	301	APENMHIDPRRKEELDKKVEDGPFMWSLSDFRQFSRLIEICNLSPDSLSSEEVHKMLVL	360
QY	361	MDENMRGSGTAGCCRNYPNTFMNPOYLTKLEEDDEDED--GESGCTFLVGLQKHRR	418
Db	361	ENGMTRGSGTAGCCQNYPATYTNPOFKIRLDEVEDQDEBSIGECCTVYLGMLQKNRW	420
QY	419	QRRMGDMHTIGFIYVEBELSGQTNHLSKNFELTNRARERSDTFLNREVLNRFCLP	478
Db	421	RKRIGMGMSIGIYAVQVPEKLESHTDHGRDFLAVQPSARTSTYVNLAEVSGRALP	480
QY	479	PGEYIILVSTEFENKQDGCIFVFSKKADYAVUDELEANLEEDDIE-DDIDGVRRL	538
Db	481	PGEYILVVPSTFEFKDGECLRFVSEKKAQALEIDVYAGNPFIEHPSEVQEDDQFRRL	540
QY	538	FAOLAGEDAIEISAFELQTIIRRYLARODIKSGDSIECTKIMVMDLSDSGKLGLKEF	598
Db	541	FELTAKDSEITRNALAKILINAFSKRTDIKQDGNINTCHEMISLDSNQTGLGAVEF	600
QY	598	YLIWTKIQYQKQIYREIDVDSGTMNSYEMRKALEEAGFKMPCQLHQYIVARFADQOLI	657
Db	601	KTILMLKIQKYLEIYMETVDYHSGTIDAHEMRLPALRKAQGTINSQYQOTIALRYVCSKGI	660
QY	658	DPDNFVRCVRLLETLEKIFKQLDPEPTGIIIELDLSWLC	696
Db	661	NFDSFVACMIRLETLEFKLSLDEDDKGNVQSLAEWLC	699
RESULT 11			
AAAB62152 standard; Protein; 739 AA.			
AAAB62152;			
29-MAY-2001 (first entry)			
Novel human protein (NHP) #1.			
Novel human protein; NHP, calcium; protease; gene therapy; screening.			
Homo sapiens.			

XX WO200116336-A1.
 XX
 XX 08-MAR-2001.
 XX
 XX 01-SEP-2000; 2000WO-US24062.
 XX
 XX 02-SEP-1999; 99US-0152057.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX
 XX WPI; 2001-235112/24.
 XX
 XX N-PSDB; AAF57302.
 XX
 XX Novel human polynucleotides isolated from human testis cDNA library,
 XX encodes novel human proteins, useful as reagents in screening for
 XX compounds used for treating mental and biological disorders -
 XX
 XX Claim 2; Page 27-28; 39pp; English.

CC The invention relates to novel human polynucleotides encoding protein
CC (NHP) that share sequence similarity with human calcium dependent
CC proteases. Nucleotide constructs encoding functional NHPs are used in
CC gene therapy for the modulation of NHP expression. NHP oligonucleotides
CC can be used as hybridization probes for screening libraries and assessing
CC NHP gene expression patterns. The sequences may also be used as part of
CC ribozyme and/or triple helix sequences that are useful for NHP gene
CC regulation. Labeled NHP nucleotide probes can be used to screen a human
CC genomic library. The NHP nucleotide sequences are also useful in drug
CC screening techniques. Nucleotide constructs encoding NHP products can be
CC used to genetically engineer host cells to express NHP products *in vivo*,
CC these genetically engineered cells function as bioreactors in the body,
CC delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion
CC protein to the body. The present sequence represents a NHP.

Sequence	739 AA
SQ	

Query Match	54.4%	Score 2020;	DB 22;	Length 739;
Best Local Similarity	52.5%	Pred. No. 7.8e-174;		
Matches 366;	Conservative 134;	Mismatches 193;	Indels 4;	Gaps 2;

```

OY 3 G1AKLAKDREAAEGLSHERALITYLNODYBALRNECLEAGTLFODPSFPAIPALGFXE 62
Dh 37 GMYHINNSRLKAGVGOHDNAQONFGNSQSEBELPAALCIKRGELFEDFLPAEPBSLGFKD 96
OY 63 LGPEYSKTRGMRMRPREICADPOFIIIGSARTPIDCGALGDDCMLAAIASITLNEEILA 122
Dh 97 LGPNSKVVQNIISWPRKDIINNPLFJIMDGISPTDICOGLIDDCCMLAAISGLTTCPLKILY 156
OY 123 RVPVLNDSPOENYAGIFHFQFWQGEWEVWVDDRILPTKGGELLFVHASGESFEWSALL 182
Dh 157 RYVRGRGSFKKNYAGIFHFQFWQGEWEVWVDDRILPTKDKILVFWHSTERSEWSALL 216
OY 183 KAVAKINGCYEALSGGATTEGFEDEFTGIAEWELKRPNNLFXIKALQKSGILGCSI 242
Dh 217 KAVKKSJYEALSGSSTMGLEDFTGVAQNSQDLQRPQNLRLAKAVERSLSMGCSI 276
OY 243 DYSADSEAITFOKLKYGHAYSYTGAEVEEVSNGSLQKLIRIRNPWEGEVETGRMNDNCP 302
Dh 277 EVTSDSELESMTDKMLVRGHAYSYTGADPHYGRKMETLIRVRNPMGRIEMNGAMSDSAR 336
OY 303 SKMNTIDEEERBLRIRHEDEBFMWSPEDEFLRHSRLTICMLTPPTLSIDPYKKKAKLRMD 362
Dh 337 EMEVEASDIDQMILAKHKTEDGEPMWSYODFLNNPLFLEICMLTPPTLSGDYKXSYNHTTFYE 396
OY 363 GMYRGSTAGCCRRNYPUTFMNNPOYLLKLEEBEDEDEDESGS---CTEVLGLIQKRRQ 419
Dh 397 GSWRGSGAGCCRRNHPTFTNRPFKLSLEBGDDPEDDABGNVYVCTCLVALMKMWRHA 456
OY 420 RKMGBDMHTIGFGLIYEVPEELSGOTNIHLKSNPELITRARERSDPTINLRVNLFXKLP 479

```

Db 457 RQGAQLOITIGFVLVAPEKQNIQDVHLKKEFTFKYODHGFSEIFITNSREVSSQLRLP 516
Qy 480 GEYILVSTPEPNKDGDFCIRVFSEKADYQAVDD-EIEMALKEPDISIDDGVRLTF 538
Db 517 GEYILVSTPEPNKDGDFCIRVFSEKADYQAVDD-EIEMALKEPDISIDDGVRLTF 576
Qy 539 AQLGDAEISAFELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGKEFY 598
Db 577 KIVAGEKEIGVVELQRLNRMALFKSFKTGFOLDACRCIMINMDKDSGKLGLEFK 636
Qy 599 ILMKIKQYKQIVKIVKIDVDRSGTNSYEMKALBEAGKRCQHLQYIVARFADQLIID 658
Db 637 ILMKIKKQWMDIFRECDHSGTNSYEMRLVIEKAGIKLNKKWQVLVARYADDDLIIID 696
Qy 659 PDNFVRCIVRLTETLTKIFKQDLPENTGTIELDLISWL 695
Db 697 DSIISCTLRKLTMTFTFLMDPKNTGHCISLEQWL 733

RESULT 12
ID AAM79025 standard; Protein: 702 AA.
AA79025;
AA79025;
AA79025;
06-NOV-2001. (first entry)
Human protein SEQ ID NO 1687.
Human; cytokine; cell proliferation; cell differentiation; gene therapy;
vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
tissue growth factor; immunomodulatory; cancer; leukaemia;
nervous system disorder; arthritis; inflammation.
Homo sapiens.
MO200157190-A2.
09-AUG-2001.
05-FEB-2001; 2001WO-US04098.
03-FEB-2000; 2000US-0496914.
27-APR-2000; 2000US-0560875.
20-JUN-2000; 2000US-0598075.
19-JUL-2000; 2000US-0620325.
01-SEP-2000; 2000US-0654936.
15-SEP-2000; 2000US-0663561.
20-OCT-2000; 2000US-0693325.
30-NOV-2000; 2000US-0728422.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Dzmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
Xue AJ, Yang Y, Wehrman T, Goodrich R,
WPI: 2001-476283/51.
N-PSDB; AAK52158.
Nucleic acids encoding polypeptides with cytokine-like activities,
useful in diagnosis and gene therapy -
Claim 20; Page 4033-4034; 6221pp; English.
The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM8020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 702 AA;
Query Match 54.2%; Score 2014; DB 22; Length 702;
Best Local Similarity 52.4%; Pred. No. 2.5e-173;
Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
Qy 4 IAAKLADREAEAGLSHERAIKYNQDYALNECEAGTLFPQDSFPALPAGKEL 63
Db 1 MVAHINNSRLKAGVGQHDNAQNGQSFELAAACRKGELFEDLPFAPESSLGKOL 60
Qy 64 GPYSKTRGWRKRPETICADPOFITGATRTDICGALGDCMLAASLTINEELAR 123
Db 61 GPNSKNQVNSWOPKDIINNPFIIMDGISPDIICGILGDCMLAAGSLTTCPKLLYR 120
Qy 124 VVPLNOSFOENYAGIFHFQFQYGEWVEYVDRLPFTKQELFVHSAGSEFWSALXK 183
Db 121 VPRGOSFKKQYAGIFHFQIWFQGVNVDRLPFTKQELFVHSAGSEFWSALXK 180
Qy 184 AVAKINGCYALSGATTEGFEDEFTGIAEWELAKPPNLFKIIQALQKSLGCSID 243
Db 181 AVAKLSGYALSGATTEGFEDEFTGIAEWELAKPPNLFKIIQALQKSLGCSIE 240
Qy 244 ITSAADSEAITPQTLKGAHAYSTGAEVENSGLQKLRIRNPQWVEVETGWNDCPS 303
Db 241 VTSDESELSWTDLVAGHAYSTGADVYRKMETLIRVNPWRIENNGAMSDSARE 300
Qy 304 WNTIDPERERLIRRHEDGSFMSFSDPLRHYSRLICNLTPTLSDTYKKKLTMDG 363
Db 301 WEEVADSIQQLHKTEDGFMSYDQFLNNFTLLEICNLTPTLSDTYKKKLTMDG 360
Qy 364 NMRGSTAGGCRNYPNTFMNPOYLITLBEDEDEDESG--CTFLVGLQKRROR 420
Db 361 SMRGSAGGCRNHPGTFWNPQKISLPEGDDPEDDAGENVVCTCLVALMKNNRHAR 420
Qy 421 KMGEDMTTIGFIYVPEELISQGTNHLKSNFLTNRARSRPTNLREVLNRFKLP 480
Db 421 QCGAQLQTLTIGFVLVAPEKQNIQDVHLKKEFTFKYODHGFSEIFITNSREVSSQLRLP 480
Qy 481 EYILVSTPEPNKDGDFCIRVFSEKADYQAVDD-EIEMALKEPDISIDDGVRLTF 539
Db 481 EYILVSTPEPNKDGDFCIRVFSEKADYQAVDD-EIEMALKEPDISIDDGVRLTF 540
Qy 540 QLAGDAEISAFELQTLIRRVLAARODIKSDGFSIETCKIMVMDLSDSGKLGKEFY 599
Db 541 IVAGEKEIGVVELQRLNRMALFKSFKTGFOLDACRCIMINMDKDSGKLGLEFK 600
Qy 600 LMTIKQYKQIVKIVKIDVDRSGTNSYEMKALBEAGKRCQHLQYIVARFADQLIID 659
Db 601 LMTIKKQWMDIFRECDHSGTNSYEMRLVIEKAGIKLNKKWQVLVARYADDDLIIID 660
Qy 660 PDNFVRCIVRLTETLTKIFKQDLPENTGTIELDLISWL 695
Db 661 DSIISCTLRKLTMTFTFLMDPKNTGHCISLEQWL 696

RESULT 13
ID AAB62154 standard; Protein: 702 AA.
AAB62154;
AAB62154;
29-MAY-2001 (first entry)
Novel human protein (NHP) #3.
Novel human protein; NHP; calcium; protease; gene therapy; screening.

XX Homo sapiens.
 OS WO200116336-A1.
 XX 08-MAR-2001.
 XX 01-SEP-2000; 2000WO-US24062.
 XX 02-SEP-1999; 99US-0152057.
 XX (LEXI-) LEXICON GENETICS INC.
 XX Donoho G, Turner CA, Nehls M, Friedrich G, Zambrowicz B, Sands AT;
 XX WPI; 2001-235112/24.
 XX N-PSDB; AAF57304.
 XX
 PT Novel human polynucleotides isolated from human testis cDNA library,
 PT encodes novel human proteins, useful as reagents in screening for
 PT compounds used for treating mental and biological disorders -
 XX
 PS Claim 6; Page 31-33; 39pp; English.
 XX
 CC The invention relates to novel human polynucleotides encoding protein
 CC (NHP) that share sequence similarity with human calcium dependant
 CC proteases. Nucleotide constructs encoding functional NHPs are used in
 CC gene therapy for the modulation of NHP expression. NHP oligonucleotides
 CC can be used as hybridization probes for screening libraries and assessing
 CC NHP gene expression patterns. The sequences may also be used as part of
 CC ribozyme and/or triple helix sequences that are useful for NHP gene
 CC regulation. Labeled NHP nucleotide probes can be used to screen a human
 CC genomic library. The NHP nucleotide sequences are also useful in drug
 CC screening techniques. Nucleotide constructs encoding NHP products can be
 CC used to genetically engineer host cells to express NHP products in vivo,
 CC these genetically engineered cells function as bioreactors in the body,
 CC delivering a continuous supply of a NHP, a NHP peptide, or a NHP fusion
 CC protein to the body. The present sequence represents a NHP.
 CC
 SQ Sequence 702 AA;
 Query Match 54.2%; Score 2014; DB 22; Length 702;
 Best Local Similarity 52.4%; Pred. No. 2.5e-173;
 Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
 QY 4 IAAKLAKDEAAGLGSHERAITYKINDYEALNECEAGTLFODSPFAIPALGFKEL 63
 DB 1 MVAHINNSRLKAKGVGHDAONFGNQSFEELRAACLRKGBELFEDPLFPAEPSSLAGFKDL 60
 QY 64 GYSSKTRGMRKRPTEICADPOFIIIGATRTDICOAGLDCWLLAIAISLTINEILAR 123
 DB 61 GPNKQVONISWQRPKQIINNPFIIMDGISPTDICQILDDCWLLAIAISLTICPKLYR 120
 QY 124 VVPLNGSFQENYAGIFHFQWQYGEWVYVVDRLPTKDGELLFVHSAEGSEFSALEK 183
 DB 121 VVREGOSFKKNYAGIFHFQWQYGEWVYVVDRLPTKDNKLVFVHSTERSSEFSALEK 180
 QY 184 AYKINGCYEALSGATTEGFEDEPTGIAEWELKKRPPLPKIOLAKGSLIGSID 243
 DB 181 AYAKLSYALSGGSTMELGDEFTGVAOSFQIQRPONLRLKRAVVERSSIMGSGISIE 240
 QY 244 ITSADSEALTFOKLVKGAHAYSVTGABEVEVNSGLQKILIRINPWGEVETGHWNDPCPS 303
 DB 241 VTSDSLESMWTKMLVKGHAYSVTGLQDVHVRGMETLIVRNPWGRIENNGAMSDARE 300
 QY 304 WNTIDPEERERLRRHEDGEFMSFSDFLHYSRLKICNTLPDTLTSDTYKKWLTVMG 363
 DB 301 WEVVASDIOQLAKTKEDGEFMSYQDFLNNFTLLKICNTLPDTLSGDYSYWHHTTYEG 360
 QY 364 NMRGSGTAGGGRNYPNTFMMNPQVLLKLEDEDEDESGSG---CTFLVGLLQKHRRRCR 420
 DB 361 SWRGSGAGGGRNHPGTFMNPQFKISLPRGDEDDAGNYYVCTCLVALMQNMHRAR 420

QY 421 KMGEDMHTTIGFIYEVPEELISGQTNHLSKNFFLTNRARERSDTFINLREVLRFKLPG 480
 DB 421 QCGAQLQTTIGFVLYAVPEKQFONIDVHLKKEFFTKYQDHFSEIFNNSREVSQRLRPG 480
 QY 481 EYTLVSTPEPNNDGPFCTRVSEKADYQAVD-ELEANLEFDSIEDIDGVRRLFA 539
 DB 481 EYTLIPSTPEPNHDAPFLKRVFEKSESWELDEVVAEQLOEKVSEDMDDPFLHFK 540
 QY 540 QLAGEDAEISAFELQTLIRVLAKRODIXSDGSIETCKIMVMDLSDSGKGLKEFYI 599
 DB 541 IVAGEGEIGVYELQRLNRMALFKFSFKTKGGLACRCMIMMDKDSGKGLKEFYI 600
 QY 600 LMTKIOKYKIYEIDVDSGTMNSYEMRKALAEAGFKQPCQLHOYIVARFADQILIDF 659
 DB 601 LMKKLKMKMDIFRECDQDSGTMNSYEMRLVIEKAGIKLNKKYQVLYVARADDDILIDF 660
 QY 660 DNFRCLVRLKTLFKTKOLDPENTGTIELDISML 695
 DB 661 DSFISCFRLKTMFTFTFLMDPNTGHICLSIQWL 696
 RESULT 14
 AAB46595
 ID AAB46595 standard; Protein; 702 AA.
 XX
 AC AAB46595;
 XX
 DT 10-APR-2001 (first entry)
 XX
 DE Human testis-specific calpain CAPN11 protein.
 XX
 KW Calpain; CAPN11; testis-specific protein; fertility regulation;
 KW calcium-dependent cysteine protease; germ cell apoptosis;
 KW testis-specific transcriptional regulator; male fertility disorder.
 XX
 OS Homo sapiens.
 OS
 PN DE19928021-A1.
 XX
 PD 21-DEC-2000.
 XX
 PF 18-JUN-1999; 99DE-1028021.
 XX
 PR 18-JUN-1999; 99DE-1028021.
 XX
 PA (BADI) BASF AG.
 XX
 DR WPI; 2001-081733/10.
 DR N-PSDB; AAF25803.
 XX
 PT New nucleic acid encoding testis-specific calpain-11, useful for
 PT identifying specific inhibitors for treatment of fertility disorders -
 XX
 PS Claim 1; Page 8-9; 16pp; German.
 XX
 CC This invention describes a novel nucleic acid (1) encoding human
 CC calpain, designated CAPN11 which has fertility regulating activity.
 CC CAPN11 is a calcium-dependent cysteine protease, expressed in testes and
 CC possibly involved in germ cell apoptosis or regulation of testis-specific
 CC transcriptional regulators. CAPN11 is useful for identifying its specific
 CC inhibitors (A). (A) are used for treating fertility disorders in men or
 CC more generally any condition associated with elevated levels of CAPN11.
 CC
 SQ Sequence 702 AA;
 Query Match 54.2%; Score 2014; DB 22; Length 702;
 Best Local Similarity 52.4%; Pred. No. 2.5e-173;
 Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2;
 QY 4 IAAKLAKDEAAGLGSHERAITYKINDYEALNECEAGTLFODSPFAIPALGFKEL 63
 DB 1 MVAHINNSRLKAKGVGHDAONFGNQSFEELRAACLRKGBELFEDPLFPAEPSSLAGFKDL 60

QY	6	GPVSKTRGMRKRPREICADPOFIIGATRRIDICQALGDCWCLAAIASLTUNLEIAR	123
Db	61	GPMSKVNQVINSWRPKDITINNPFLFIDGISTPDIQGIIGDCLAAIGSLTTPCKLYR	1200
QY	124	VVPLNQSFCQNTYVAGIFPHFQFMOYGEVWEVVDRLRPTDQDELLEFVSAEGSEFWASLLEK	1833
Db	121	VVRGQSFKKNVAGIFHFPQIMQCGQVNVNVVDRLRPTNDKDLVFHSHRESEFWASLLEK	1800
QY	184	AVAKINGCYAALSGGATTBGFEDFTGGIAEWELKPKPPNLFKIIQKALQKSGSLGCSID	2433
Db	181	AVAKLSGSSYALSGGSGTMEGLBDFTGVAQSGFOLQRPQNLRLRKRAKVAERSISMGCSIE	2400
QY	244	ITSAOSEALTPOKLVKGHAYSTGSAEVEENSQSLQKLRIRNPMGEVEMTGRNNDNCPSS	3033
Db	241	ITSDSELESTMDKLVRGHAYSTGLQDHYRGKMETLIRVNPWGRLEMGASDSARE	3000
QY	304	WNTIDPEERERLTRRHEDGEFWMSFSDFLRHYSRLEICNLTPDTLSDTYKKMLTKMDG	3633
Db	301	WEEVASDIQWQLHKTEDEGFWMSSYDFLNNLTLEICNLPTDLTSGYKSYWHTTYEG	3600
QY	364	NMRGSGTAGGCNNYENITFMNPOYLKLEEDDEDEDESGS--CTPLVGLIQKRRRQR	4200
Db	361	SMRGSASAGSCRNPGTFWNPQFKISLPEGDPEDDAEGNVVCTCLVALMQKMRHAR	4200
QY	421	KMGEDHTTGFGIYEVPEELSGQTNHLSKSNFLNRAERSDFTINLRVLPFKLPBG	4800
Db	421	QOGAQLOTTGIFVLYAAPKEQNIQDHLKKEFFTYQDHGSELIPTNRSEVSSQLRLPG	4800
QY	481	EYLLVPTPEPNKDGPFCTRFVESEKADYQAVDD--EIEANLEPDISDDIDDDGRLLFA	5333
Db	481	EYLLIPSTPEPHDADFLRVPTFKGSEMEDEVNVYAEQLQEEVSEDDVDQDFLHLFK	5400
QY	540	QLAGDEAIEISAFELQTLRLRVLAQRDIIKSDFSIETCKINWMLDSDGSKGLKPEFYI	5993
Db	541	IVAGEKEIGVYELQRLNRMALFKFSSPTKFGFDACRCMINLMDKSGSKLGLTEFKI	6000
QY	600	LWTKIQKQKQYKIVEIVDNRSGTNSVEMKALKEEGFPMQCHQVYARADQULIIDF	6553
Db	601	LWKLKMKWMDIFRECDQDHSGLTNSYEMKLVEKGKIKLNKMWQVLVARYADDLIIIDF	6600
QY	660	DNEFVRCVLRLETFLFKIKOLDPENTSTIELDLISWL	695
Db	661	DSFISCTLRKIMFTFFLMDPNKNTIHICLSLEQNL	696
RESULT 15			
AAU72887	ID	AAU72887 standard; Protein; 702 AA.	
XX	XX	AAU72887;	
AC	XX		
XX	XX		
DE	DT	26-FEB-2002 (first entry)	
XX	XX		
XX	XX	Human asparyl protease partial protein sequence #12.	
KM	XX	Human; protease; PCR primer; cytosstatic; immunomodulatory; cardiant;	
KM	XX	vasootropic; antimigraine; analgesic; endocrine; nootropic; tranquiliser;	
KM	XX	hyperensitive; hypotensive; neuroleptic; neuroprotective; anabolic;	
KM	XX	anorectic; antiinflammatory; asparyl protease; cysteine protease;	
KM	XX	metalloprotease; serine protease; cancer; hematopoietic; breast; colon;	
KM	XX	lung; prostate; cervical; brain; ovarian; bladder; kidney; pain;	
KM	XX	immune-related disease; cardiovascular disease; neuronal disease;	
KM	XX	immune; sexual dysfunction; mood disorder; attention disorder;	
KM	XX	cognition disorder; hypotension; hypertension; psychotic disorder;	
KM	XX	dyskinesia; metabolic disorder; inflammatory disorder.	
OS	XX	Homo sapiens.	
XX	XX	WO200183782-A2.	
XX	XX	08-NOV-2001.	
PD	XX		
PF	XX	04-MAY-2001; 2001WO-US14431.	

XX	04-MAY-2000; 2000US-201879P.
PR	(SUGC-) SUGEN INC.
PA	Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
XX	Payne V;
PI	WPI; 2002-041502/05.
XX	N-PBDB; AAS97170.
DR	
XX	
PT	Novel protease polypeptide useful for screening for substances that may
PT	be used to treat, e.g., cancers, immune-related diseases,
PT	cardiovascular disease, migraine, pain, psychotic and inflammatory
PT	disorders -
PS	Claim 28; Figure 2C; 232pp; English.
XX	
CC	The invention relates to an isolated, enriched, or purified protease
CC	polypeptide (I) and polynucleotide (II) encoding (I). (I) may be used to
CC	screen for substances (S) that may modulate its activity. Administering
CC	S (which modulates protease activity in vitro) may be used to treat a
CC	disease or disorder selected from cancers (e.g., of tissues, of blood or
CC	haematopoietic origin, of the breast, colon, lung, prostate, cervical,
CC	brain, ovarian, bladder or kidney), immune-related diseases and
CC	disorders, cardiovascular disease, brain or neuronal-associated diseases
CC	(e.g., central or peripheral nervous system diseases, migraine, pain,
CC	sexual dysfunction, mood disorders, attention disorders, cognition
CC	disorders, hypotension, hypertension, psychotic disorders, neurological
CC	disorders and dyskinesias), metabolic disorders and inflammatory
CC	disorders. (I) may also be useful as a diagnostic tool for a disease or
CC	disorder such as those above. AAU72876-AAU72910 represent human
CC	protease amino acid sequences of the invention.
CC	
XX	
SQ	Sequence 702 AA;
Query Match	54.2%; Score 2014; DB 23; Length 702;
Best Local Similarity	52.4%; Pred. No. 2.5e-173;
Matches 365; Conservative 134; Mismatches 193; Indels 4; Gaps 2	
QY	4 IAAKLAKREARAEGLSHERAIKYUNQVYLARNCELAGTLFODSPFPALSGFKEL 63
DB	1 MVAHINNRLAKKGVGHNDNQNPONGSFEEELRAACLAKGELFEDPLFPARESSLGFDL 60
QY	64 GPYSKTRGMWRKPTEICADPOFIIGATRTDIOGALGDCLLAIASLTINEILAR 123
DB	61 GPSNKNVNISMQRKDIIINPLIFIMDISPTDICOGLIGDCWLAAIGSLTTCCKLYR 120
QY	124 VVPLAQSTQENVYAIGTFHFQMQYGEHWVVVDRLPTDGELLFPHSAGSEFWALLDK 183
DB	121 VVPRQSFSKKNVAGIFHHFOIQFGQWVAVDDRLPTINDKLVFHSERSEFWALLDK 180
QY	184 AYAKINGCYEALSGGATGEPDFPGIAEWYELKKPPNLFKILOKLOSLSLGCSD 243
DB	181 AYAKISGYEALSGGSTTWEGLEDFFGVGAOSFOLRRPONLRLLKRAVERSLNGCSIE 240
QY	244 ITSADSEAITFOKLVKHAYSVTGAEEVESNGSLÖKLIIRLNPFGEVMTGRANDNCPS 303
DB	241 VTSDBELSESMTDKMLVRGHAYSVTGLDVHRHGKMETLIRVFNPMGRILEMNGAMDSARE 300
QY	304 WNTIDPEERREKULTRRHDEGFEMMSFDPRLRIHYSLEICNLTPDTLTSPTYKKMKULTKMDG 363
DB	301 WEVVASDIQOMOLHTKTEDGEFMWSYODLPANFTLLEICNLTPDTLSGDYKSYMHTTFYEG 360
QY	364 NMRGSSTGGCRNYPNTEFMANNPOVYLKLEEDDEDDEDESG---CTPIVLGIÖKRBROR 420
DB	361 SWRRSSXAGGCRNHGFTWPNQPKTSLPEGDDDDDAEGNVVCTCLVALAQKWRRHAR 420
QY	421 KMGDMTHITIGIYEVEPEELSQTQHILSKNFPLTNRARERSDFTINTREVLRNKFLPG 480
DB	421 QOGAQLQGTTIGVLVAVPKEPQNIDVHLKKEFTTYQDHGSEIFTNSREVSQGLRPBG 480
QY	481 EYLVIPSTFEENKQDGFICIVFSEKKADYQAVDD-ELEANIEEDFISBEDDIDGVRRLFA 539

Db 481 EYIIPTFEPHRDADFLRLRVTEKHSWELDEVNTAEOLOEKEVSEDDMODFLHLFX 540
 QY 540 QLAGEDAEISAFELQTLIRVLAERODIKSDGFSIETCKIWDMLDSDGSGKLGKEFYI 599
 Db 541 IVAGEGEKEIGVYELQRLNMAIKFKSGFKTGFGIDACRCMINIMDKDGSGLLEPKI 600
 QY 600 LWTXIQYQKIYREIDVDRSGTMNSYEMRKALBEAGFMPCOLHOVIVARPADDLIIDF 659
 Db 601 LMKLKKWMDIFRECDQDHSGSTLNSYEMRLVIERKAGIKLNKVMQVLVARYADDDLIIDF 660
 QY 660 DNFVRCIVRLLETFKIKOJDPENTGTIELDLISML 695
 Db 661 DSFISCFRLKTMFTFPLTMDPKNTGHICLSLEOML 696

Search completed: July 24, 2003, 12:51:44
 Job time : 91.9793 secs

